

# 会話によるニュース記事伝達のためのトリビアの獲得と活用

## Acquisition and Utilization of Trivia for Conversational News Contents Delivery

高津 弘明<sup>\*1</sup> 松山 洋一<sup>\*1</sup> 本田 裕<sup>\*2</sup> 藤江 真也<sup>\*1\*3</sup> 小林 哲則<sup>\*1</sup>  
Hiroaki Takatsu Yoichi Matsuyama Hiroshi Honda Shinya Fujie Tetsunori Kobayashi

<sup>\*1</sup>早稲田大学 Waseda University <sup>\*2</sup>本田技術研究所 Honda R&D Co.,Ltd. <sup>\*3</sup>千葉工業大学 Chiba Institute of Technology

Trivia has characteristics that draw people's interests and has been used as a strategy for attracting interests in contents in various conversational applications, such as entertainment and educational sites. In this paper, we propose a model that distinguishes trivia based on the surface representation of sentences. We incorporated the trivia presentation function into the spoken dialogue system we are developing and assessed the effect of presenting trivia in the introduction part of the topics in the news transmission task. As a result of the experiment, we confirmed that it is effective for enhancing users' interests in topics and promoting the feeling of rapport with system although we were not able to find significant improvements in the number of users' feedback and the Efficiency of Information Transfer (EoIT).

### 1. はじめに

文の表層表現に基づいてトリビアを識別するモデルを提案する。また、ニュースを伝える会話システムにおいて、話題の導入部分でトリビアを提示することの効果を検証する。

我々はニュース記事のようなまとまった量の情報を効率的に伝達する会話システムの開発を行っている [高津 18]。ここで「効率的」とは、伝達対象となる記事の中から、ユーザにとって不要な情報を除き、必要な情報だけを伝えることを意味する。我々のシステムの特徴は、あらかじめ主計画、副計画と呼ぶ複数のシナリオを用意しておき、このシナリオに沿って会話を進めることで、リズムの良い会話を実現するうえで必須となる迅速な応答を可能としたところにある。主計画に沿って記事の要点となる情報を提示する傍らで随時ユーザからのフィードバックを理解し、必要に応じて副計画に遷移して補足情報を提示する。このようにユーザの興味や理解状態に応じて提示する情報を柔軟に切り替えながら会話を進めていく仕組みを持つ。一方で、高い情報伝達効率 (EoIT; Efficiency of Information Transfer) を実現するにはユーザからのフィードバックが必要不可欠である。本研究ではフィードバックに寄与する要素としてユーザのコンテンツに対する興味とシステムに対する親密感に着目する。

トリビアには人の興味を引く性質があり、従来よりコンテンツに対する興味を誘引する戦略として、会話主体のエンターテインメント (ゲーム、クイズショーなど) や教育現場などでトリビアが利用されてきた [Brown 84, Palincsar 84, ProMarketing Wizard 13, Voices Heard Media 13]。

人の興味を引くということに関して、落語の世界では聴衆を「話を聞かせる」状況に置くために、話の「マクラ」を導入部分に持つてくることが多い [広瀬 10]。ここでマクラは、次に来る話題への説明的な導入の機能を果たし、話し手と聴衆の間の親密感 (ラポール) の構築に寄与していると考えられる。

本研究では、話題の導入部分でのトリビアの提示がコンテンツに対する興味の誘引だけでなく、システムに対する親密感 (ラポール) の増進にも効果があると仮定する。興味や親密感が改善されることで、ユーザのフィードバックが増加し、情報伝達効率や継続利用可能性の向上が見込めるか、実験を通して検証する。

### 2. 関連研究

#### 2.1 トリビアの自動獲得

トリビアには、人の興味を引く性質があり、ゲームなどのエンターテインメントや教育現場などで利用されてきた [Brown 84, Palincsar 84, ProMarketing Wizard 13, Voices Heard Media 13]。例えば、Brown と Palincsar らは、読解力の指導においてトリビアは学生の注意をコンテンツに向けさせる効果があると指摘している [Brown 84, Palincsar 84]。

トリビアを自動で獲得する手法に関する研究やトリビアを対話システムの発話生成に応用した研究がある。新名らは、ユーザを対話に引き込む効果を期待し、料理についての雑学を発話する対話システムを開発した [新名 17]。また、トリビアを含む文を「ある物事についての瑣末な知識であり、かつ人の興味を引くような内容を含む文」と定義し、主語の IDF や分散表現、主語と文中の名詞との類似度などを素性として文のトリビア度合いを回帰やランキング学習によって推定する手法を提案した [新名 18]。太田らは、対話を盛り上げるような発話の生成を目的として、Wikipedia の各文を TF-IDF、語の共起、文の長さによってスコア付けし、意外性のある文の抽出を行った [太田 09]。Prakash らは、固有表現の種類や Superlative Wordsなどを素性として Rank SVM を学習し、トリビアらしさのランキングを行った [Prakash 15]。Tsirel らは、トリビア度を人物とカテゴリの類似度の低さから計算される surprise とカテゴリに含まれる全ページ間の平均類似度の高さから計算される cohesiveness の積で定義し、人物の Wikipedia ページに付与されたカテゴリからトリビア度を計算する手法を提案した [Tsirel 17]。Fatma らは、(entity, predicate, object) の三つ組からなるデータを入力として、CNN でトリビアかどうかを識別する手法を提案した [Fatma 17]。また、人手で設計した特徴量と組み合わせることにより高い精度でトリビアを識別できることを示した。

本研究でも、従来トリビアと呼ばれてきたものと同様、トリビアを「ある物事についての瑣末な知識であり、興味をそそるもの」と定義する。また、Fatma らと同様にトリビアか否かを識別する 2 値分類問題として定式化するが、本研究では、文の単語系列そのものを入力として、双方向 LSTM と self-attention [Lin 17] からなるモデルによって識別を行う。そして、ニュース伝達タスクにおいて、トリビアの提示がユーザのコンテンツに対する興味を誘引する効果があると仮定し検証する。

連絡先: 早稲田大学 理工学術院 知覚情報システム研究室  
〒162-0042 東京都新宿区早稲田町 27  
E-mail: takatsu@pcl.cs.waseda.ac.jp

## 2.2 会話のマクラと話し手ー聞き手間のラポール形成

人間同士のインタラクションを円滑にし、協働的なタスクの達成率を向上させるためには、対話相手同士の繋がりや親密さの感覚「ラポール (rapport)」が十分に形成されていることが重要であると、社会心理学や人工知能などの分野で指摘されてきた。Spencer-Oatey はラポールを対話の調和度合いと円滑さであると定義した [Spencer-Oatey 05]。特に、交渉 [Zhao 18], カウンセリング [Kang 12], 教育 [Zhao 14] など、相互の親密度が特定のタスクの達成に貢献するようなドメインでラポールの効果が報告されている。近年、ラポールの機能を部分的に実装した会話エージェントが提案されている。Pecune らは、情報推薦型のバーチャルアシスタントにラポール機能を導入し、ラポールの度合いと情報の受諾率などについて検討している [Pecune 18]。

ところで、「聴衆に話を聞かせるメディア」であるところの落語は、伝統的に「マクラ」、「本題」、「落ち」を基本構造として持つ。一般的に、落語家はいきなり落語の演目に入突することはほとんどなく、聴衆に語りかける雰囲気をつくるために、まずマクラとして、挨拶をしたり、世間話をしたり、軽い小咄を披露したりしてから本題に入っていく。マクラは本題への導入部であり、話し手と聴衆の心理的距離感を縮め、本題に関連する話題で聴衆の意識を物語の現場に引きつけるような効果があるとされている [広瀬 10]。このように、効果的に「話を聞かせる」状況をつくるために、マクラは話し手と聴衆の間のラポールの構築に寄与していると考えられる。

以上のことから、本研究では冒頭でのトリビアの提示がユーザのシステムに対する親密感を増進する効果があると仮定し検証する。

## 3. トリビアの識別

双方向 LSTM と self-attention [Lin 17] からなるモデルによってトリビアか否かを識別する手法を提案する。モデルの全体像を図 1 に示す。

### 3.1 トリビア識別モデル

識別対象の文が  $n$  個の単語系列  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$  から構成されているとする。  $\mathbf{x}_i$  は  $i$  番目の単語の one-hot ベクトルである。まず、各単語を埋め込み層で  $d$  次元に圧縮し、得られた単語の埋め込み表現を双方向 LSTM に入力する。ここで、LSTM の隠れ層の次元を  $u$  とする。

$$\mathbf{e}_i = W_{xe} \cdot \mathbf{x}_i \quad (1)$$

$$\vec{\mathbf{h}}_i = \text{LSTM}(\vec{W}_{eh} \cdot \mathbf{e}_i + \vec{W}_{hh} \cdot \vec{\mathbf{h}}_{i-1}) \quad (2)$$

$$\overleftarrow{\mathbf{h}}_i = \text{LSTM}(\overleftarrow{W}_{eh} \cdot \mathbf{e}_i + \overleftarrow{W}_{hh} \cdot \overleftarrow{\mathbf{h}}_{i+1}) \quad (3)$$

順方向 LSTM の出力と逆方向 LSTM の出力を連結し、それらを  $n$  単語分並べて得られる行列を  $H$  とする。

$$\mathbf{h}_i = [\vec{\mathbf{h}}_i || \overleftarrow{\mathbf{h}}_i] \quad (4)$$

$$H = (\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n) \quad (5)$$

この行列  $H$  に対して attention を計算するが、長い文において複数箇所 ( $r$  箇所) に注目できるように  $r \times d_a$  次元の行列  $W_{s2}$  を掛ける ( $d_a$  はハイパーパラメータ)。

$$A = \text{softmax}(W_{s2} \tanh(W_{s1} H^T)) \quad (6)$$

得られた attention 行列  $A$  を入力文の隠れ層の行列  $H$  と掛けあわせて文の埋め込み行列  $S$  を得る。この  $r \times 2u$  次元の行列  $S$  を  $2ru$  次元のベクトル  $\mathbf{s}$  に変換し、文の埋め込みベクトルを得る。

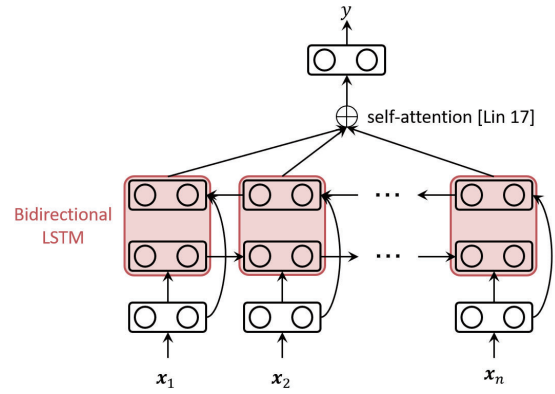


図 1: トリビア識別モデル

表 1: トリビアデータセットの統計

|    | 訓練セット | 開発セット | テストセット |
|----|-------|-------|--------|
| 正例 | 2500  | 100   | 126    |
| 負例 | 2500  | 100   | 126    |

表 2: トリビアと非トリビアの例

| トリビア                                       | 非トリビア                        |
|--|------------------------------|
| 童話「ウサギとカメ」で 2 匹が競争した距離は 2 8 0 m 3 3 c m 未満 | 最初は馬の姿で現れるが、命じられれば人間の姿になる    |
| アルゼンチンには、昔の丸の内線が走っている                      | また、自動車関連メーカーにはより安全な自動車の開発を促す |
| ミツコビナマケモノは首を最大 2 7 0 度回転させられる              | ネオジム塩を成分とする魷物の記事も含ます         |

$$S = AH \quad (7)$$

$$\mathbf{s} = [S[1] || S[2] || \dots || S[r]] \quad (8)$$

最後にこのベクトルをシグモイド関数を活性化関数とする出力層に与え、トリビアである確率を計算する。

$$y = \sigma(W_{sy} \mathbf{s} + b) \quad (9)$$

### 3.2 評価実験

Web から収集したトリビアを正例、Wikipedia からランダムに抽出してきた文を負例としてデータセットを作成し、トリビア識別モデルの性能を評価する。

#### 3.2.1 実験設定

Web から収集したトリビアを正例、Wikipedia からランダムに抽出してきた文を負例 (非トリビア) とする。負例を抽出するにあたり、トリビアらしい文が含まれないように負例は次の 3 つの条件を満たす文の中からランダムに抽出した。(1) 最初の段落でない、(2) 段落の先頭の文でない、(3) 記事タイトルを含まない。作成したトリビアデータセットの統計を表 1 に示す。また、トリビアおよび非トリビアの例を表 2 に示す。

モデルのネットワーク構造として、埋め込み層の次元  $d$  を 100、LSTM の隠れ層の次元  $u$  を 100、attention の次元  $d_a$  を 100、attention の数  $r$  を 5 に設定した。訓練セットにおいて頻度が 2 以下の単語は特殊記号 (unk) に置き換え、入力の語彙サイズを 4790 に制限した。

表 3: トリビアと非トリビアの識別精度 (%)

| Accuracy | 「トリビア」<br>識別精度 | 「非トリビア」<br>識別精度 |
|----------|----------------|-----------------|
| 97.2     | 97.6           | 96.8            |

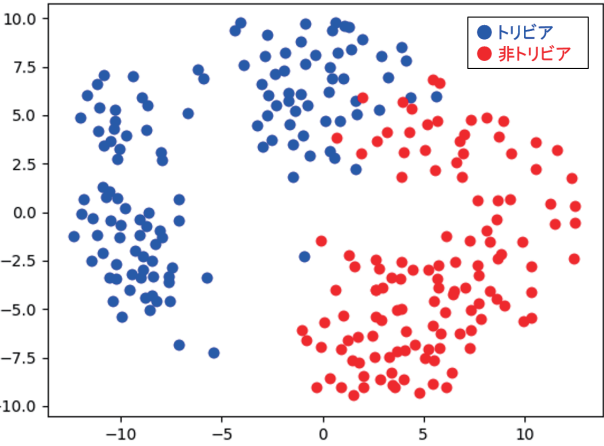


図 2: テストセットにおけるトリビアと非トリビアの埋め込み表現  $s$  の分布を多次元尺度構成法で次元圧縮可視化した結果

3.2.2 実験結果

テストセットにおけるトリビアと非トリビアの識別精度を 3 に示す。この結果から、提案モデルによって 97%と高い精度でトリビアか否かを識別できることがわかった。

テストセットにおけるトリビアと非トリビアの埋め込み表現  $s$  の分布を多次元尺度構成法で次元圧縮可視化した結果を図 2 に示す。この結果から、トリビアと非トリビアの境界がある程度明確に分かれるように学習できていることがわかった。

3.3 Wikipedia からのトリビアの獲得

学習した識別モデルを Wikipedia の記事に適用しトリビアを獲得する。ここでは、文脈依存性の高い文を回避するために以下の条件を満たす文に対して識別を行った。(1) 最初の段落である、(2) 段落の先頭の文である、(3) 記事タイトル+「は」を含む文である。Wikipedia の約 100 万記事に適用し、約 10 万文のトリビアを獲得した。

4. トリビアの会話システムへの導入効果の検証

ニュースを伝える会話システム [高津 18] においてトリビアを提示する効果を検証する。話題の導入部分でその話題に関連したトリビアを提示することで、ユーザのコンテンツに対する興味やシステムに対する親密感が向上すると仮定する。さらに、それによって、ユーザのフィードバックが増加し、情報伝達効率や継続利用可能性が向上するか、実験を通して検証する。

4.1 仮説

- 以下の仮説について検証を行う。
- 仮説 1: トリビアがトピックへの興味の増進に寄与する
  - 仮説 2: トリビアがラポール形成に寄与する
  - 仮説 3: トリビアがフィードバック数の増加に寄与する
  - 仮説 4: トリビアが情報伝達効率の向上に寄与する
  - 仮説 5: トリビアが継続利用可能性の向上に寄与する

表 4: 冒頭でのトリビア提示の例

| 発話タイプ | 発話内容   |
|-------|--|
| トリビア  | そういえば、天気予報の 177 って、昔は大隈重信の自宅の電話番号だったらいいよ                           |
| 話題遷移  | それで、その天気予報に関するニュースなんだけど  |
| 主計画   | 6 月から天気予報が大きく変わるみたいだよ<br>気象庁が計算能力を約 10 倍に高めたスーパーコンピュータを新たに導入するんだって |
| ...   | ...  |

表 5: アンケート項目

| ID | 質問内容                             |
|----|----------------------------------|
| I1 | ～の話について、対話しているうちに興味が強まった         |
| C1 | 自分とシステムが同調していたと感じた               |
| C2 | システムと対話しているとき、あまり心地よくないことが言えなかった |
| C3 | 対話の間、フラストレーションがたまった              |
| M1 | 対話システムが自分の言っていることに注意を向けてると感じた    |
| M2 | 対話システムがユーザに対して敬意を払っていると感じた       |
| P1 | 対話システムがフレンドリーと感じた                |
| P2 | 対話システムに温かみを感じた                   |
| P3 | 対話システムは自分を気遣ってくれていると感じた          |
| R1 | このシステムをまた使用したい                   |

4.2 比較条件

- 以下の 2 つの条件を比較する。
- 条件 1: ベースライン (主計画+副計画)
  - 条件 2: 条件 1 + トリビア提示機能
- 条件 2 における会話冒頭のトリビア提示の例を表 4 に示す。なお、条件 1 では冒頭の 2 行がなく主計画から会話が始まる。

4.3 実験設定

各条件ごとに、本システムを始めて利用する大学生 8 人ずつ、3 つのニューストピック (平均 33.7 文、主計画サイズ 9 文) について本システムと会話させ、会話終了後にアンケート調査を行った。

アンケート項目を表 5 に示す。I1 は各トピックについての興味に関する評価項目である。R1 は継続利用可能性に関する評価項目である。残りがラポールに関する評価項目である。

ラポールに関しては、Tickle-Degnen らが分類した同調性 (Coordination)、相互注意 (Mutual Attentiveness)、ポジティブ性 (Positivity) という 3 つの要素に従いアンケート項目を作成した [Tickle-Degnen 90]。ここで、C1,C2,C3 が同調性、M1,M2 が相互注意、P1,P2,P3 がポジティブ性に関する評価項目である。

各評価項目について 5 段階で評価させた (5 点: そう思う, 4 点: どちらかと言えばそう思う, 3 点: どちらでもない, 2 点: どちらかと言えばそう思わない, 1 点: そう思わない)。

また、被験者には、会話終了後にニュース記事を読ませ、各文について会話で伝えて欲しかったかどうかを 5 段階で評価させた (5 点: 伝えてほしい, 4 点: どちらかと言えば伝えてほしい, 3 点: どちらでもよい, 2 点: どちらかと言えばいらない, 1 点: いらない)。会話で提示できた 4 点以上の文の被覆率と会話で提示しなかった 2 点以下の文の除外率の調和平均から情報伝達効率を被験者ごとに計算した。



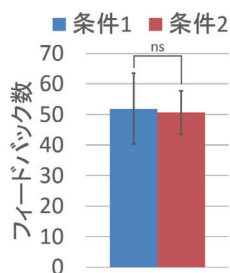


図 3: フィードバック数

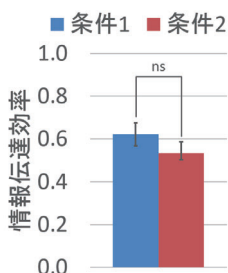


図 4: 情報伝達効率

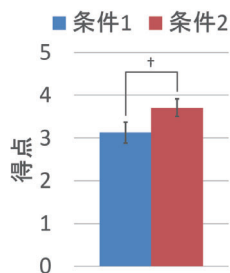
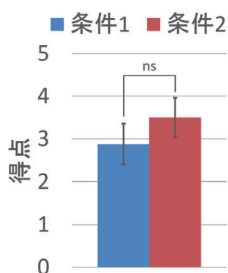
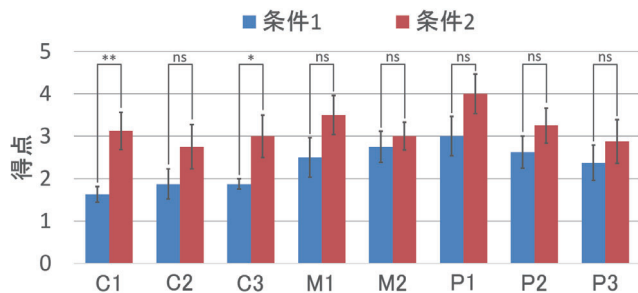
図 5: 興味増進効果 (†:  $p < 0.1$ )

図 6: 継続利用可能性

図 7: ラポール形成効果 (\*:  $p < 0.05$ , \*\*:  $p < 0.01$ )

#### 4.4 実験結果

各条件について平均値を計算し、t検定による仮説検定を行った。被験者一人あたりの平均フィードバック数を図3に示す。被験者一人あたりの平均情報伝達効率を図4に示す。興味増進効果(I1)と継続利用可能性(R1)に関するアンケートの平均点を図5、図6に示す。ラポールに関する各評価項目の平均点を図7に示す。なお、C2とC3は同調的でないことを質問しているため、図は同調的であるほど値が高くなるように6から引いた値を示している。

フィードバック数と情報伝達効率、継続利用可能性については有意な差が得られず、仮説3と仮説4、仮説5の正しさは示されなかった。一方、興味増進効果とラポール形成効果(特に、同調性)については有意な差が得られ、仮説1と仮説2の正しさが示された。

#### 5. おわりに

トリビアか否かを識別するモデルを提案し、97%と高い精度で識別できることを確認した。本研究では、文の表層表現のみを入力としてトリビアの識別を行ったが、先行研究で検討されているようなトリビアらしさを表現する特徴量も加えることでさらなる性能向上が見込めると考えられる。

検証実験の結果、情報伝達を目的とした会話システムにおいて、トリビアを話のマクラとして冒頭で伝えることにより、ユーザのトピックに対する興味とシステムに対する同調感が増進されることがわかった。すなわち「話を聞かせる会話メディア」

という新しいジャンルにおいて、トリビアの戦略的な使用が聞き手を効果的に「聞くモード」に置かせる方法として有効であることが示唆された。

しかしながら、本来の目的であった、高い情報伝達効率(EoIT; Efficiency of Information Transfer)を実現するためのフィードバックの促進効果については確認されなかった。

ユーザを「聞くモード」にしたうえで、他にどのような因子が直接的にフィードバックの促進に寄与するのかは、未だ解っていない。今回の検証実験の自由記述アンケートでは、そもそも期待した質問応答が得られなかった、音声合成の質が良くないと感じられる、間(ま)が悪いなどの、システム性能に対する信頼を疑う意見も散見された。今後の研究では、システム性能やラポール現象の網羅的な測定と分析を通してフィードバックを促進させるメカニズムをより系統的に明らかにしていく。

#### 参考文献

- [Brown 84] Brown, A.L., Palincsar, A.S., and Armbruster, B.B.: Instructing comprehension-fostering activities in interactive learning situations, *Learning and Comprehension of Text*, pp. 255-286 (1984)
- [Fatma 17] Fatma, N., Chinnakotla, M.K., and Shrivastava, M.: The unusual suspects: Deep learning based mining of interesting entity trivia from knowledge graphs, in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 1107-1113 (2017)
- [広瀬 10] 広瀬和生: 現代落語の基礎知識, 集英社, ISBN 978-4-08-771365-7 (2010)
- [Kang 12] Kang, S.-H., Gratch, J., Sidner, C., Artstein, R., Huang, L., and Morency, L.-P.: Towards building a virtual counselor: Modeling non-verbal behavior during intimate self-disclosure, in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 63-70 (2012)
- [Lin 17] Lin, Z., Feng, M., Santos, C.N., Yu, M., Xiang, B., Zhou, B., and Bengio, Y.: A structured self-attentive sentence embedding, in *Proceedings of the 5th International Conference on Learning Representations*, pp. 1-15 (2017)
- [新名 17] 新名和也, 嶋田和孝: シズルワードから想起される料理の雑学を話す対話システム, 電子情報通信学会信学技報, Vol. 117, No. 367, NLC2017-42, pp. 77-82 (2017)
- [新名 18] 新名和也, 嶋田和孝: トリビア文抽出のためのトリビア度合いの推定, 電子情報通信学会信学技報, Vol. 118, No. 122, NLC2018-7, pp. 69-74 (2018)
- [大田 09] 太田知宏, 島海不二夫, 石井健一郎: 発話生成を目的とした Wikipedia からの文抽出, 第 23 回人工知能学会全国大会論文集, 2G1-NFC5-11, pp. 1-4 (2009)
- [Palincsar 84] Palincsar, A.S. and Brown, A.L.: Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities, *Cognition and Instruction* 1, No. 2, pp. 117-175 (1984)
- [Pecune 18] Pecune, F., Chen, J., Matsuyama, Y., and Cassell, J.: Field trial analysis of socially aware robot assistant, in *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 1241-1249 (2018)
- [Prakash 15] Prakash, A., Chinnakotla, M.K., Patel, D., and Garg, P.: Did you know? - Mining interesting trivia for entities from wikipedia, in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, pp. 3164-3170 (2015)
- [ProMarketing Wizard 13] ProMarketing Wizard: Trivia game on facebook, <http://bit.ly/1HUhlRx> (2013)
- [Spencer-Oatey 05] Spencer-Oatey, H.: (Im)politeness, face and perceptions of rapport: Unpackaging their bases and interrelationships, *Journal of Politeness Research. Language, Behaviour, Culture*, Vol. 1, No. 1, pp. 95-119 (2005)
- [高津 18] 高津弘明, 福岡維新, 藤江真也, 林良彦, 小林哲則: 意図性の異なる多様な情報行動を可能とする音声対話システム, 人工知能学会論文誌, Vol. 33, No. 1 (2018)
- [Tickle-Degnen 90] Tickle-Degnen, L. and Rosenthal, R.: The nature of rapport and its nonverbal correlates, *Psychological Inquiry*, Vol. 1, No. 4, pp. 285-293 (1990)
- [Tsurel 17] Tsurel, D., Pelleg, D., Guy, I., and Shahaf, D.: Fun facts: Automatic trivia fact extraction from wikipedia, in *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 345-354 (2017)
- [Voices Heard Media 13] Voices Heard Media: Using trivia and quiz products to engage your customer, <http://bit.ly/1Htiv5V> (2013)
- [Zhao 14] Zhao, R., Papangelis, A., and Cassell, J.: Towards a dyadic computational model of rapport management for human-virtual agent interaction, in *Proceedings of the International Conference on Intelligent Virtual Agents*, pp. 514-527 (2014)
- [Zhao 18] Zhao, R., Romero, O.J., and Rudnicky, A.: SOGO: A social intelligent negotiation dialogue system, in *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pp. 239-246 (2018)