



HAL
open science

Text Data Analysis on Answers Written in Japanese to Free Text Questions Obtained at Astronomy Lectures

Seiichiro Aoki, Kazushi Sakka, Keiji Emi, Shinzo Kobayashi, Toshio Okamoto

► **To cite this version:**

Seiichiro Aoki, Kazushi Sakka, Keiji Emi, Shinzo Kobayashi, Toshio Okamoto. Text Data Analysis on Answers Written in Japanese to Free Text Questions Obtained at Astronomy Lectures. Open Conference on Computers in Education (OCCE), Jan 2020, Mumbai, India. pp.139-143, 10.1007/978-3-030-59847-1_15 . hal-03519209

HAL Id: hal-03519209

<https://hal.inria.fr/hal-03519209>

Submitted on 10 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

Text Data Analysis on Answers written in Japanese to Free Text Questions obtained at Astronomy Lectures

Seiichiro Aoki^{1,2} [0000-0001-6391-8781], Kazushi Sakka¹, Keiji Emi¹, Shinzo Kobayashi¹ and Toshio Okamoto¹

¹ The Kyoto College of Graduate Studies for Informatics, 6068225, Japan

² Kyoto University, 6078471, Japan

s_aoki@kcg.ac.jp

Abstract. Questionnaires were completed at the astronomy lectures “AstroTalk”, to improve promotion of “AstroTalks”, by identifying satisfaction of participants, based on statistical analysis of answers. Textual data analysis was performed using KH Coder on answers of free text questions written in Japanese. In this paper, we show results of the analysis of the questions: “What were the impressive things in our lecture?” and “What topics of astronomy are you interested in?” Strong connections among words and terms related to topics of “AstroTalks” are shown in co-occurrence network diagrams of the answers to the these questions. We adopted topics in “AstroTalks” based on results, to improve lecture design. In addition, we are developing an active learning system with an electronic blackboard, connecting lectures and participants through the Internet to reduce regional disparities in educational opportunities in astronomy.

Keywords: Text Data Analysis, Free Text Questions, KH Coder.

1 Introduction

As science communication activities, we have been offering astronomy lectures and tours since 2011. The tour is an astronomical walking tour with historical features entitled “Millennium Trail of Astronomy in Kyoto” [1]. The lecture’s title is “Millennium Trail of Astronomy in Kyoto AstroTalk” - in short, “AstroTalk”. It consists of an ordinary astronomy lecture and a special astronomy lecture with stereoscopic videos named “Kyoto4D” and the software “Mitaka”. The special lecture is called “4D Space Theater”, where participants can feel the three-dimensional structure of the sun, planets, stars and galaxies, and the vastness of the universe. “Kyoto4D” is a set of stereoscopic videos produced by the Astronomical Observatories of Kyoto University, and “Mitaka” is provided by the 4D2U project of NAOJ. “AstroTalk” has been held 4-5 times a year since 2011, usually at Kyoto University Museum (a total of 38 times so far). In every “AstroTalk”, we distribute questionnaire for participants to get statistical data in order to improve the promotion of our activity and satisfaction of participants. The questionnaire consists of multiple-choice questions and free text questions.

In this paper, we show the result of the text analysis applied to answers written in Japanese by participants to free text questions in the questionnaire (see Fig. 1(a)).

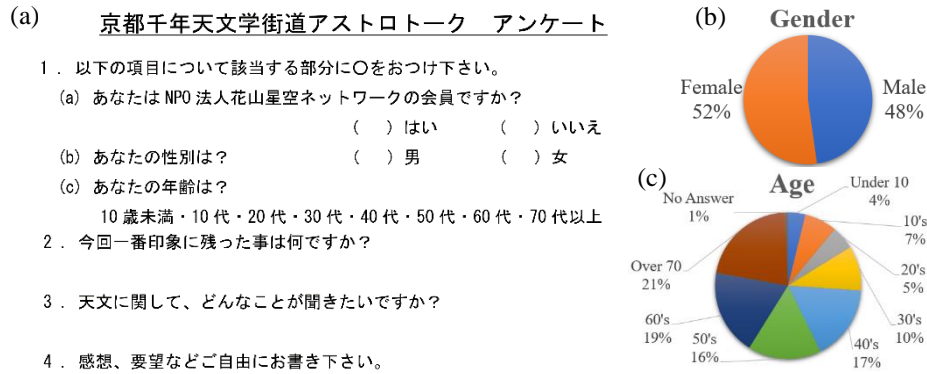


Fig. 1. (a) Shows parts of the questionnaire sheet related to this paper. Questions 2, 3 and 4 are free text questions. (b) Shows gender and (c) age distributions of participants.

2 Features of the Data

We analyzed the data obtained at “AstroTalks” from July 2011 to May 2019. The number of participants was 1,168 and 740 participants responded to the questionnaire, so the response rate was about 74.5%. The number of male and female participants was almost equal (see Fig. 1(b)). More than half of the participants were over the age of 50 years (see Fig. 1(c)). In this paper, we show the result of statistical data analysis with KH Coder [2] on the answers written by participants in Japanese to free text questions 2 “What were the impressive things on our lecture?” (556 participants answered) and 3 “What topics of astronomy are you interested in?” (353 participants answered) [3]. KH coder is widely used in various research for analysing text data [4].

3 Text Data Analysis of the Answers to Free Text Questions

As a pre-process of the analysis, we unified words or terms that had the same meaning in the answers. In fact, “4D Theater”, “4D Space Theater” and “Theater” were unified to “4D Space Theater”. Then, we performed statistical text data analysis on the answers to the free text questions using KH Coder. Table 1 shows the top 16 words and terms by the number of appearances, neglecting particles and auxiliary verbs in the answers to question 2 “What were the impressive things on our lecture?” The data showed which words or terms were had impressions on participants. In Table 1, “4D Space Theater” appeared 229 times. Thus, many participants had a strong impression of the “4D Space Theater”, which is the special lecture in “AstroTalk”. In addition, the “Sun” (21 times) and “Mars” (20 times) appeared many times in the

answers. These were often topics in the “AstroTalk”. Although the “Moon” was a topic in “AstroTalk” only 3 times, it appeared 17 times in the answers. This is probably due to the fact that the Moon is explained in every “4D Space Theater”.

Table 1. Top 16 words or terms with number of appearances.

Word or term	Number of appearances	Word or term	Number of appearances
4次元宇宙シアター (4D Space Theater)	229	知る (know)	24
宇宙 (universe)	83	太陽 (Sun)	21
映像 (image)	39	大きい (big/large)	21
地球 (The Earth)	36	超新星爆発 (Supernova Explosion)	21
星 (Star)	35	火星 (Mars)	20
見る (watch)	33	面白い (interesting)	18
話 (topic)	30	お話 (talk)	17
銀河 (galaxy)	29	月 (Moon)	17

3.1 Result of the Analysis of the Answers to a Free Text Question

Fig. 2 and Fig. 3 are co-occurrence network diagrams of answers to the questions. We can see which words and terms are strongly connected with the words and terms based on how frequently they appeared in the sentences in the answers.

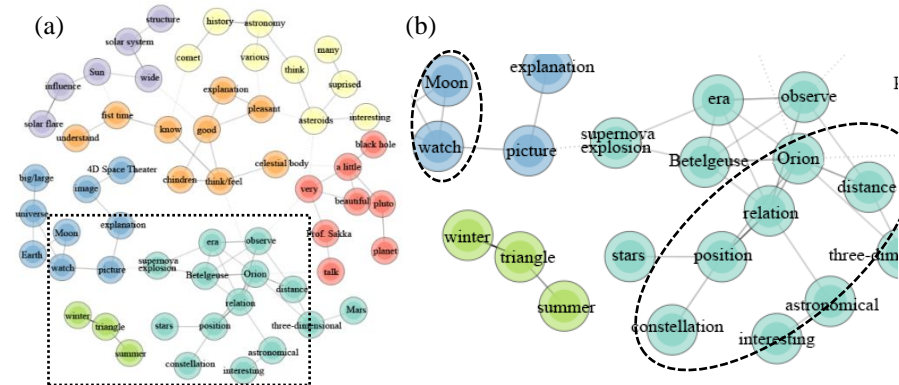
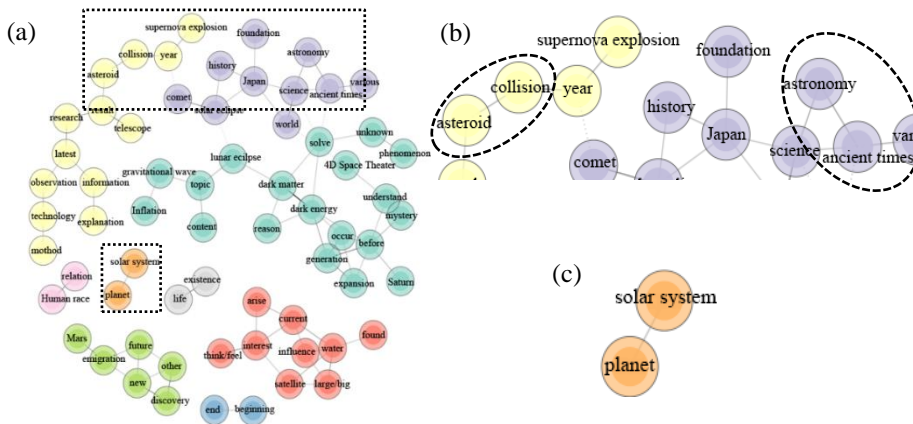


Fig. 2. (a) The co-occurrence network diagram of the answers to the free text question “What were the impressive things in our lecture?” (b) The enlarged view of the area inside the dotted rectangle in Fig. 2a. The minimum number of appearances is 5 and the maximum is 60. In each diagram, the lines indicate strong connections among words or terms.

As for the answers to question 2 “What were the impressive things on our lecture?”, the “Moon” has a strong connection with “watch” (see the upper left dotted ellipse in Fig. 2(b)). In fact, there were answers by participants such as “I was able to watch the back side of the Moon”, “I watched the video to show how the Moon was formed”.

So, we can say that the simulation video of the “Moon” formation and the back side of the Moon were strongly impressed on participants as the lecturer intended. On the other hand, connections among “Orion”, “constellation”, “distance” and “position” are also strong (see the lower right dotted ellipse in Fig. 2(b)). It is probably because the three-dimensional structure of the Orion constellation and the position of the stars and distance among the stars that form the Orion constellation were shown by the lecturer at almost every “AstroTalk”. In fact, there are answers such as “the positions of stars that form the Orion constellation” and “the distance among the stars that form the Orion constellation”. Text data analysis has identified major trends in the text data statistically, but it is not easy to find minor trends such as misconceptions.

As for the answers to question 3 “What topics of astronomy are you interested in?”, it can be seen from the diagram that the popular set of words and terms for participants were, for example, “astronomy” in “ancient times”, “collision” of “asteroid” to the Earth and “planets” of “solar system” (see dotted ellipses in Figs. 3(b) and 3(c)). In order to improve the design of “AstroTalk” based on the results of the analysis, we adopted several sets of words and terms seen in the diagram as topics for “Astro-



Talks”.

Fig. 3. (a) The co-occurrence network diagram of answers to the free text question “What topics of astronomy are you interested in?” (b) The enlarged view of the inside of the dotted rectangle at the top of Fig. 3(a), and (c) that at the bottom of Fig. 3(a). The minimum number of appearances is 3 and the maximum is 60. In each diagram, the lines indicate strong connections among words or terms.

4 Summary and Future Research

We performed statistical text data analysis using KH Coder on answers written in Japanese to free text questions which were obtained from participants at the astronomy lectures “AstroTalks”. Words and terms related to topics of “AstroTalks” appeared many times in the answers. From co-occurrence network diagrams, it was shown quantitatively that participants were impressed with the topics which the lec-

turers wanted to convey in “AstroTalks”. We adopted sets of words and terms seen in the diagram as topics for “AstroTalks” to improve lecture design.

There are regional disparities in educational opportunities in astronomy in Japan, because lectures and seminars are usually held in cities. This is because lecturers are usually professors of colleges and universities located in cities. To reduce disparities, applying the system shown, which was very effective in the education of agriculture skills [5], we are developing an active learning system on astronomy with an electronic blackboard connecting a lecturer in a city with participants at a town through the Internet (Fig. 4). Through the system, participants can easily access learning contents through a large size touch panel for group work connecting with



lecturers online.

Fig. 4. Active learning system on astronomy under development. The white box on the left side is a projector and the touch device on the right side is a 34-inch size touch panel.

Acknowledgements. This work was supported by JSPS KAKENHI (Grants-in-Aid for Scientific Research) Grant Numbers 19H01724. We thank Ms. Mami Umemoto for support in data aggregation.

References

1. Aoki, S.: “Millennium Trail of Astronomy in Kyoto” Outreach Activity: an Astronomical Walking Tour with Historical Features and Lectures. In: *Proceedings of Communicating Astronomy with the Public Conference (NAOJ)*, pp. 220-221. (2018).
2. Higuchi, K., Komoda, N., Tamura, S., Ikkai, Y.: A Support Tool for Composing Social Survey Questionnaires by Automatically Summarizing Questionnaires Stored in Data Archives. *WSEAS Transactions on Information Science & Applications* 4(2), 280-287 (2007).
3. Aoki, S.: Text Analysis of Questionnaires in Astronomical Lectures. *IPSJ Symposium Series 2019*, pp. 277-282 (2019).
4. Palmer, S., Campbell, M.: Text analytics visualisation of Course Experience Questionnaire student comment data in science and technology. In: *Australasian Association for Engineering Education Conference*, Geelong, VIC (2015).
5. Emi, K., Kobayashi, S.: Practice of Blended Learning of Next Generation Agricultural Human Resources Development Using Wearable Cameras and Drone Cameras - The “WAZA” Method. Presentation at IFIP TC3 OCCE 2018, Linz, Austria (2018).