*ICIC Express Letters*
*Part B: Applications*
Volume **12**, Number **12**, December **2021**

ICIC International ©2021 ISSN 2185-2766
pp. 1129–1136

# A STUDY ON THE EVALUATION SUPPORT METHOD FOR HUMAN RESOURCES BY FOCUSING ON THE CONTENT OF CONVERSATION IN GROUP DISCUSSIONS IN RECRUITMENT INTERVIEWS

Asako Ohno

Faculty of Engineering
Osaka Sangyo University
3-1-1 Nakagaito, Daito, Osaka 574-8530, Japan
ohno@eic.osaka-sandai.ac.jp

Received May 2021; accepted July 2021

Abstract. *Due to the COVID-19 pandemic since 2020, many companies are conducting recruitment interviews for new graduates online. In face-to-face group discussions, not only the content of the candidates' conversations, but also visual information such as eye contact, posture, and gestures is subject to evaluation. In online group discussions, however, the external information that can be obtained from examinees through the camera is limited. Calculating indicators of speaker characteristics from the content of examinee conversations and presenting them to evaluators as a basis for their decisions in evaluating examinee performance, can be expected to be beneficial to both examinee and evaluator. This paper describes the basic idea and concrete procedure of an attempt to convert the conversational content of examinees participating in a group discussion into text and to evaluate the speakers according to the extracted keywords and reports the results of a simple evaluation using a tool that implements this method.*
**Keywords:** Fundamental competencies for working persons, Online recruitment interview, Group discussion, Text analysis, TF-IDF, Intelligent information

1. **Introduction.** The COVID-19 pandemic since 2020 has forced us to change our lifestyles throughout our society. One example is the recruitment interview conducted by companies as part of their recruitment activities. The interviews conducted by Japanese companies for new graduates often include group discussions. In group discussions, candidates are tested on their communication skills and their ability to immediately find their own role in the group and how they can contribute to achieving the goal. In Japanese companies, the ability to work cooperatively with others is highly valued.

In the past, face-to-face assessments were based as much on the content of the candidate's conversation as on visual information such as eye contact, posture, and gestures. However, because of the COVID-19 pandemic, recruitment activities have been shifted to online [1], and the visual information that interviewers can obtain from candidates is replaced by two-dimensional images viewed through a camera, and the amount of information is much smaller than in face-to-face interviews.

The use of online discussions is rapidly increasing not only in recruitment activities at companies but also in university classes. For example, the results of a qualitative evaluation based on the content of online discussions conducted in university general education classes suggest that discussions among students can lead to deeper learning if clear guidelines are provided and a reasonable amount of time is spent [2]. However, judging the depth of learning from the content of a discussion requires specialized knowledge, and it is even more difficult to judge the learning depth in online discussions, which have limited

DOI: 10.24507/icicelb.12.12.1129

visual information such as gestures and facial expressions compared to face-to-face discussions.

As a result of the rapid shift to online activities due to the COVID-19 pandemic, there is a need for knowledge that is useful for making judgments that have been made by human hands based on knowledge and experience in face-to-face situations. If a quantitative objective index that supports human judgment is presented to the evaluator, it is expected to be easier to realize. In our previous study [3], we reported a group discussion speaker evaluation method that aims to support active learning by estimating the depth of learning using keywords extracted from the content of speakers' conversations based on the definition of the Bloom's Revised Taxonomy [4].

In this study, we focus on group discussions in online interviews. We partially modified the group discussion speaker evaluation method for active learning support proposed in our previous study and attempted to apply it to the evaluation in recruitment interviews. Specifically, keywords are extracted from the text of group discussion conversations for each examinee and assigned to the three evaluation items of "ability to step forward", "ability to work in a team", and "ability to think things through", which are the "fundamental competencies for working persons" [5] proposed by the Ministry of Economy, Trade and Industry (METI) and widely used in recruitment interviews. See [6] for an explanation of "fundamental competencies" for working persons in English. The results of the sorting are visualized, and the characteristics of the examinees obtained from the content of the conversations are presented as objective data for personality evaluation, with the aim of supporting the personality evaluation conducted by human resource managers.

The contribution of the study is to support the judgment of human resource managers in recruitment interviews by extracting and quantitatively presenting the characteristics of speakers from the content of discussion conversations based on interviews with corporate human resource managers and university employment support staff and METI's "fundamental competencies for working persons". The rest of this paper is organized as follows. Section 2 explains the background of the research, and Section 3 explains the idea and the procedure of speaker analysis in this research. Section 4 reports the results and findings of applying the method to an actual online group discussion, and Section 5 concludes the paper.

## 2. Background.

### 2.1. Recruitment of new graduates by Japanese companies in the FY2020 COVID-19 pandemic.

Due to COVID-19, hiring activity in 2020 is slower than in previous years. Mynavi Inc. conducted a survey of companies that hired new graduates in June 2020 [1]. Of the 2,565 companies that responded, about 36.6% hired 0% of new graduates, while about 22.4% hired between 10% and 40%. The percentage of companies that hired less than 10% has increased by 9.1 percentage points from the previous year (fiscal 2019), when companies struggled to hire in what is known as a "seller's market".

### 2.2. Related works.

Text mining approaches have been used to analyze a variety of applications. Lee et al. [7] collected keywords from 11,052 patents on AI registered in Korea and used text mining techniques to analyze the relationship between AI technology and the development requirements for collaborative robots that help in the efficient operation of smart factories. They defined derived keywords for each of the 11 development requirements, counted the occurrences of the keywords in patent documents related to various AI technologies, and conducted a factor analysis. Their method and our method have in common that they use a list of the frequency of occurrence of keywords defined for each category. However, in this method, the share of each category shown in the list is directly used as a classification table to keep the idea of applying the Revised Taxonomy in our previous method and to make the results interpretable.

Although there is a limited amount of existing research on text analysis related to recruitment activities, the following is a brief overview of the research. Asano et al. [8] conducted a simple text mining of free text in reports on jobs for which the recommended applicants were not hired at a public institution called "Hello Work" to analyze the factors of failure. The results for each age group showed that the word "interview" was in the top 10 most frequently used words for all 10, 20, and 30s age groups, but not for 40s, 50s, and 60s (while the opposite trend was seen for the word "age").

In recent years, with the spread of active learning for the purpose of realizing proactive learning, discussion analysis in the classroom has been actively conducted even before COVID-19. However, most of them are based on face-to-face discussions, and they rely more on visual and audio information than on what is said. Naim et al. [9] built a model and analyzed 138 videos of 69 university students practicing interview with career counselors. The model used variety of information: 1) visual information such as smiles and head gestures, 2) textural information such as word counts and topic modeling, and 3) prosodic information such as pitch, intonation, pauses of the interviewees. With that information, they estimate the three evaluation criteria: excitement, friendliness, and engagement. From the findings, they recommended to speak more fluently, use less filler words, speak as "we" instead of "I", use more unique words, and smile more.

This result suggests that the content of speech, i.e., textual information, plays an important role as useful information for evaluation in recruitment interviews. Although it is desirable to use a variety of information as in this study, it is also necessary to emphasize the acquisition of information from text in online interviews, because images and sound information are scarcer in online interviews than in face-to-face interviews.

Several AI-based corporate interviewing services have already been commercialized [10-12]. Many of them use both external information such as the examinee's gaze obtained from camera images and the text content of the speech. As an evaluation method using the content of the speaker's speech, there are some that evaluate and present the results of interactive questioning with the candidate based on a predetermined scenario, and others that estimate the candidate's psychological state by performing sentiment analysis on the text of the speech.

No attempt has been made yet to present an objective index based on METI's fundamental competencies for working persons. Although this method is intended to support human resource managers as a stand-alone tool, it can be used in combination with existing tools to bring the candidate information obtained from online interviews closer to that of face-to-face interviews.

2.3. **Evaluation indicators for group discussions in recruitment interviews.** The points that recruiters often check in group discussions during recruitment interviews include "aggressiveness", "cooperativeness", "communication skills", "ability to think logically", and "ability to come up with ideas". Group discussions in recruitment interviews have some common evaluation indicators.

As a result of an interview survey with three human resource managers from companies listed on the first section of the Tokyo Stock Exchange, we found that although there are certain evaluation indicators in group discussions during recruitment interviews, subjective evaluations are inevitably involved because the evaluations are made by the people who are at the interview. In addition, online recruitment interviews can be conducted without time and place constraints using interactive AI. AI interviews are expected to reduce the burden and cost associated with human interviews, as well as to standardize hiring criteria and eliminate inequalities in subjective evaluations by people.

On the other hand, however, there is also a persistent expectation for the "eyes" of human resource managers, given that they are selecting "people" who will work with them for many years. The interview survey also emphasized the importance of the "personality aspect" rather than just current abilities in hiring new graduates. These findings indicate that while a common objective index is necessary to reduce the burden and cost of recruitment interviews and to make fair judgments, subjective judgments by human resource managers are also supported.

## 3. Proposing Method.

**3.1. Basic idea.** In this study, we propose a person evaluation support method focusing on online group discussions, which are rapidly increasing in the Corona disaster. As can be seen in Figure 1, in conventional group discussions, not only the content of the conversation but also visual information such as movements, posture, and speaking style can be obtained. However, in online group discussions, the visual information is much less than in face-to-face discussions. In this study, we propose a person evaluation support method that extracts and evaluates keywords from the content of conversations where the information obtained is the same as in face-to-face discussions.

In this study, we propose a method to support personality evaluation by extracting keywords from conversations that are not different from those obtained in face-to-face situations. By placing importance on the contents of conversations during group discussions, it becomes possible to select examinees on more equal terms. In addition, we would like this research to be a steppingstone to a system that automatically evaluates everything from group discussions to presentations in the future.

This method follows the previously proposed speaker evaluation method [3] for group discussions to support active learning, with some modifications. In the previously proposed method, keywords extracted from the conversational content of group discussions are counted to determine which of the six categories of the cognitive process dimension of the Revised Taxonomy [4] known as the "taxonomy of educational objectives": "remember", "understand", "apply", "analyze", "evaluate", and "create", they fall into. In this way, the learner's learning process can be assessed. This supports the implementation of active learning by quantitatively presenting students' learning depth to teachers.

The Revised Taxonomy clearly categorizes educational goals by placing specific educational goals on a two-dimensional table called the "cognitive process dimension" and the "knowledge dimension". The "cognitive process dimension" contains a list of "taxonomy action verbs", which are specific actions that learners are expected to take once they have reached one of the six categories that make up the "cognitive process dimension" (i.e., they have achieved the educational goal associated with one of the categories).

Using extended version of taxonomy action verbs, the proposed method estimates which cognitive process dimension a learner corresponds to based on the keywords in the discussion conversation that indicate the actions a learner is taking.

**3.2. TF-IDF.** In this method, TF-IDF (Term Frequency-Inverse Document Frequency) is used to extract characteristic keywords for each speaker from the text of the group discussion. TF-IDF is used to extract important words that are characteristic of a particular document (discussion) and/or unnecessary words that appear in all documents.

TF-IDF is known as one of the preprocessing methods in natural language processing. It is used to investigate the features of documents based on the importance of words in the documents, to calculate the similarity between documents, and to perform document retrieval. TF (Term Frequency) is the number of occurrences of a word $t_i$ relative to $\sum_{t_k \in d_j} f(t_k, d_j)$, the sum of the number of occurrences of any words $t_k$ in a document $d_j$ as shown in Equation (1). The higher the frequency of a word, the larger its TF value.

$$tf(t_i, d_j) = \frac{f(t_i, d_j)}{\sum_{t_k \in d_j} f(t_k, d_j)} \tag{1}$$

IDF (Inverse Document Frequency) is the rarity of a word. IDF is the number of documents containing a word $t_i$ against the total number of documents $N$, and the rarer a word is, the higher its value is.

$$idf(t_i) = \log\left(\frac{N}{df(t_i)} + 1\right) \tag{2}$$

The right-hand side of Equation (2) is logarithmic to reduce the effect of the number of documents. To prevent *idf* from becoming zero, 1 is added. TF-IDF is the multiplication of the TF value and the IDF value as shown in Equation (3).

$$tf \cdot idf(t_i, d_j) = tf(t_i, d_j) \cdot idf(t_i) \tag{3}$$

3.3. **Evaluation categories and keywords table.** The previously proposed method quantitatively expressed the students' learning depth by classifying and counting the keywords extracted from the discussion using the "taxonomy classification table" created based on the definition of the Revised Taxonomy.

In this method, instead of Revised Taxonomy, we adopt the "fundamental competencies for working persons", which are proposed by METI and used as evaluation criteria in employment interviews and consist of the three categories of "ability to step forward", "ability to work in a team", and "ability to think things through". Specifically, keywords are extracted from the text of group discussion conversations, and when a keyword matches one of the three categories in the "fundamental competencies for working persons classification table", it is assigned to that category, and the number of matches is counted.

To construct a classification table of discussion keywords based on the "fundamental competencies for working persons", we first summarized the specific behaviors of examinees who possessed the three "fundamental competencies for working persons" in Table 1, based on the contents of job-hunting support books and web articles.

Table 1. Specific actions that correspond to the three evaluation categories

| Evaluation criteria | Ability to step forward | Ability to work in a team | Ability to think things through |
|---|---|---|---|
| Specific actions | · Actively express one's opinion.<br>· After expressing agreement with others' opinions, give an additional opinion. | · Elicit the opinions of others.<br>· Listen to others' opinions while giving them feedback.<br>· Try not to get sidetracked from the discussion. | · Cite objective evidence and explain things using numbers and data.<br>· Make brief comments. |

The keywords that are associated with specific behaviors for each of the evaluation items in Table 1 are defined and arranged in Table 2, which is called the "fundamental competencies for working persons classification table (FCC table)". The procedure for creating the classification table is as follows.

First, the keywords associated with the evaluation items were selected from the contents of conversations in group discussions and books on recruitment interview preparation. After that, we interviewed three human resource managers of companies listed on the first section of the Tokyo Stock Exchange to examine their vocabulary and used a thesaurus to obtain synonyms of the examined vocabulary. Finally, the validity of the vocabulary was confirmed with the help of a counselor at the career center of Osaka Sangyo University.

Table 2. Fundamental competencies for working persons classification table (FCC table)

| Evaluation criteria | Keywords |
|---|---|
| Ability to step forward | 役割, 補足, 補う, 足す, プラス, 補充, 補完, 付け足す 加える, 増補, まず, 第一に, 前者, 後者, 思う<br>(Translation: role, supplement, fill, add, plus, replenish, complement, add to, add, augment, first, first of all, former, latter, think) |
| Ability to work in a team | いかがですか？, どう思いますか？, ありますか？, 共有, 整理, 分担, 時間配分, ○分, シェア, 認識, どのように, いつ, どこで, なにを, なぜ, どうして, 教えて下さい, はい, ええ, なるほど, うん, やっぱり, もっともです, 本当ですか, そうですね, 賛成, 同じ, その通りです, 説明, 解説, 明らかにする, 明白にする<br>(Translation: How about it? How do you think? Do you have any? share, organize, role-sharing, time allotment, $n$ minutes, share, recognize, how, when, where, what, why, why not, tell me, yes, yes, I see, yes, yes, that's right, is that true? yes, it is., I agree, same, yes, you're right, explain, describe, reveal, clarify) |
| Ability to think things through | 具体的, データ, 規則性, 効率, ベースにして, もとに, 詳しく, 丁寧に, それで, そこで, そして, それから, だから, 順を追って, 筋道立てて, 噛み砕いて, 一方で, 分かりやすく, まとめる, 合わせる, 統一, 要するに, 定義, 前提, ○つ, 理由, キモ, 重要, ○%, ○点, 他にも, 例えば, 以外, その他, あとは<br>(Translation: concrete, data, regularity, efficiency, based on, based on, detailed, careful, then, there, and then, so, in order, logical, crammed, on the other hand, easy to understand, summarize, match, unify, in short, definition, premise, $n$ items, reason, key, important, $n$%, $n$ points, other, for instance, except, and so on, what else) |

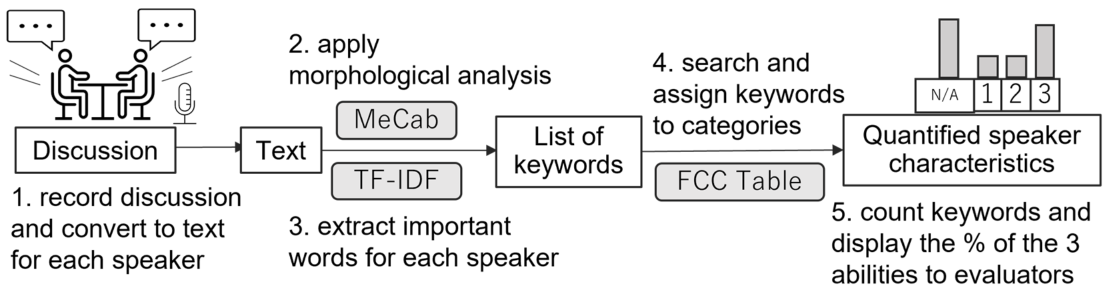


Figure 1. Procedure of the method

3.4. **Procedures.** Figure 1 shows the procedure of our proposed method. First, we preprocess the recorded text of the group discussion by eliminating unnecessary words and performing morphological analysis. Next, TF-IDF is used to obtain a set of characteristic keywords for each speaker. The result of matching and counting the keywords of each evaluation item in the FCC table shown in Table 2 is presented to the evaluator as the result of the speaker's personality evaluation.

## 4. Evaluation.

4.1. **Settings.** For analysis, we use a discussion video distributed by a channel for job-hunting students on *YouTube* [13]. The topic of the discussion was "How can we raise Japan's food self-sufficiency rate?" The participants in the discussion were five university students who were engaged in job hunting. The group discussion lasted 20 minutes. Since the faces of the examinees are not shown in this video, the visual information obtained in a face-to-face interview is insufficient.

4.2. **Results and findings.** We prototyped a tool that implements this method using python and conducted speaker analysis of text-based discussions. Table 3 shows the results of matching the keywords extracted from the speech content of each speaker with the keywords of each evaluation item in the FCC table, and Figure 2 shows the evaluation results presented to the evaluator.

TABLE 3. Number of keywords counts for each evaluation category in the FCC table

| Categories \ Speakers | A | B | C | D | E |
|---|---|---|---|---|---|
| Ability to step forward | 1 | 3 | 1 | 1 | 1 |
| Ability to work in a team | 1 | 2 | 1 | 3 | 4 |
| Ability to think things through | 5 | 4 | 5 | 2 | 3 |
| N/A | 139 | 126 | 138 | 57 | 123 |

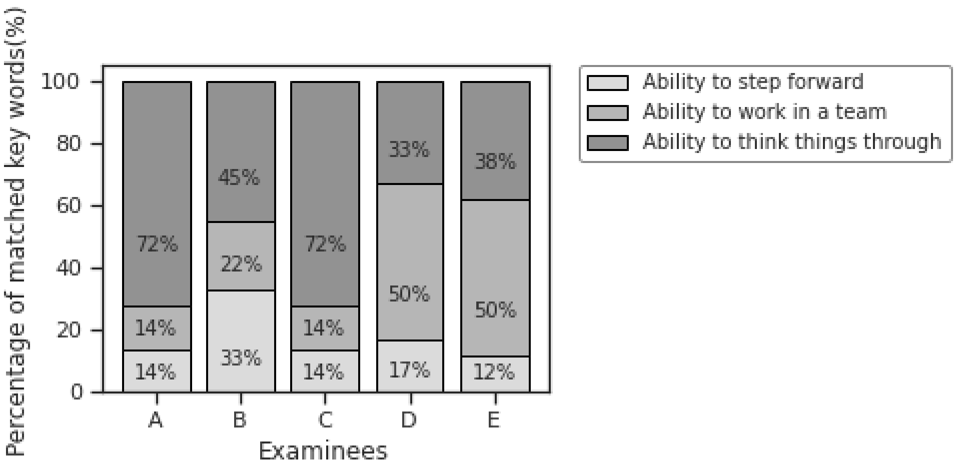


FIGURE 2. Percentage of the number of keywords extracted from the speech of each of the five examinees (A to E) that correspond to the three evaluation categories (Only the value with the largest first decimal place will be rounded up, all other values will be rounded off)

As shown in Table 3, since there are few keywords defined in the FCC table and only exact matches are counted for keywords at present, there are few keywords assigned to each category of evaluation items, and the majority are assigned to "N/A". Synonyms, fluctuations, and colloquial expressions need to be addressed.

Although the number of extracted keywords was small, the results in Figure 2 are consistent with the evaluation conducted visually by watching the video. For example, A has the strongest impression of "ability to think things through" in the visual evaluation because he often digs into the missing viewpoints and summarizes opinions expressed by others. The following is a concrete example of such an exchange.

C: *"But agriculture is completely different depending on what you are producing."*

A: *"Does it have anything to do with the weather?"*

Here is another example. The impression of "ability to work as a team" is the strongest in the visual evaluation of D, because D says less and gives more responses to others' opinions.

E: *"I think you're too focused on that."*

D: *"Well, since there are only seven minutes left after the three points, I would like to hear your opinion on the first point, PR for improving the image of farmers."*

5. **Conclusions.** The rapid increase in online recruitment interviews since COVID-19 has limited the visual and audio information that is used to play a major role in face-to-face recruitment interviews, leaving human resource managers with insufficient information to assess the characteristics of candidates for group discussions. A survey conducted by a private company in FY2020 highlighted the stagnation of hiring activities and the current difficulty of implementing online interviews in many small and medium-sized companies.

In this study, we applied our previous study to quantitatively evaluating group discussion speakers in recruitment interviews based on the "fundamental competencies for working persons" of METI. We have confirmed that the framework of our method can provide useful quantitative information for evaluators in a simple procedure.

The method is still in its early stages and needs to be improved. For example, the number of keywords associated with each evaluation item in the FCC table is small. In our previous method, we acquired synonyms using word2vec. Furthermore, notation fluctuation and context need to be considered.

## REFERENCES

[1] Mynavi Corporation, *Survey of Corporate Recruiting Activities for College Graduates in 2021*, https://saponet.mynavi.jp/news/news_file/file/column-570_001.pdf, Accessed on October 14, 2021.

[2] S. A. Aderibigbe, Can online discussions facilitate deep learning for students in general education?, *Heliyon*, vol.7, no.3, pp.1-6, 2020.

[3] A. Ohno, A study on quantitative representation of student learning depth in group discussions using a revised taxonomy, *Proc. of the 6th International Conference on Electronics and Software Science*, pp.31-37, 2020.

[4] L. W. Anderson et al., *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*, Longman, New York, 2000.

[5] Ministry of Economy Trade and Industry (METI), *Winners Selected for the 2016 Grand Prix for Development of Fundamental Competencies for Working Persons*!, https://www.meti.go.jp/english/press/2016/0223_01.html, Accessed on October 14, 2021.

[6] T. Ioi, Development of fundamental competencies for working persons through project-based learning, *Management Studies*, vol.2, no.3, pp.159-167, 2014.

[7] H. J. Lee, J. S. Kim and H. W. Kim, Analysis of artificial intelligence technology based on the requirements of collaborative robots through patent analysis, *ICIC Express Letters*, vol.13, no.6, pp.521-527, 2019.

[8] H. Asano et al., Study on hiring decision: Analyzing rejected applicants by mining individual job placement data of public employment security offices, *Procedia Computer Science*, vol.60, pp.1156-1163, 2015.

[9] I. Naim et al., Automated prediction and analysis of recruitment interview performance: The role of what you say and how you say it, *Proc. of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pp.1-6, 2015.

[10] Talent and Assessment Inc., *Strategic Hiring AI Navigator (SHaiN)*, https://shain-ai.jp/, Accessed on October 14, 2021.

[11] Web Spiders Inc., *The Leading AI-Based Video Interview Bot – Gecko*, https://gecko.ai/, Accessed on October 14, 2021.

[12] Mya Systems Inc., *The AI Recruiting Assistant Most Loved by Candidates – Mya*, https://www.mya.com/, Accessed on October 14, 2021.

[13] Live Group Discussion, *DEiBA Job Hunting Channel*, https://www.youtube.com/watch?v=0nEcl6cSOB8, Accessed on October 14, 2021.