

# Prototype of Decision Support Based on Estimation of Group Status Using Conversation Analysis

Susumu Kono<sup>1(✉)</sup> and Kenro Aihara<sup>1,2</sup>

<sup>1</sup> SOKENDAI (The Graduate University for Advanced Studies),  
Tokyo, Japan

{su-kono,kenro.aihara}@nii.ac.jp

<sup>2</sup> National Institute of Informatics,  
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

**Abstract.** We propose a prototype for a system to verify the decision support of a group based on estimations of group status through utterance analysis.

Based on methods used in prior studies of group dynamics and utterance analysis, we measured the utterance characteristics of group members to infer group status; moreover, we aimed to enhance the overall condition of the group by providing appropriate reference information in a timely manner through a conversational agent system. The goal of this system was a more satisfying decision-making process.

We manufactured a prototype system to verify both the operations involved in the test case and the ability to infer group classification and status according to group dynamics.

Our proposed method was appropriate based on the prototype system. Future work will focus on optimizing the logic and system functions of group status estimation and the subsequent step of informing the group.

**Keywords:** Conversation estimation · Group status estimation · Utterance feature · Conversational agent · Intention extraction

## 1 Introduction

### 1.1 Object

Our objective is to clarify the effectiveness of estimating the classification and status of a group by measuring the utterance characteristics of the group members. In particular, we aim to explore the possibility of identifying all utterances in a conversation using the extracted intention of some words in utterances and time course information at the first step.

Then, we will explore the possibility of estimating connections of utterances and relationships between participants in the conversation using this information and replacing any missing information with utterance characteristics. The relationships between participants that are revealed will be applied for estimating the classification and status of a group at the next step.

## 1.2 Background and Motivations

In group decision-making, some members may hesitate or yield their position to superiors, or they may offer no concrete ideas or suggestions, and thus these members may feel frustration in such situations. Group members who are in a weak position may not be as confident, and thus cannot clearly declare their intentions, while members who hold a position of power in the group can confidently and clearly declare their intentions.

Therefore, our motivation for this study was to solve this problem by using speech recognition and the extraction of intentions based on a group dynamics approach. We also aim to lead the overall group to a good situation by providing appropriate reference information and suggestions in a timely manner using status estimations based on a group dynamics approach.

## 2 Related Work

### 2.1 Group Dynamics

We define a *group* as an aggregate of individuals who have frequent interaction, mutual influence, common feelings of camaraderie, and work together to achieve a common goal. We define a *member* as an individual who joins a group.

Group dynamics refers to a system of behaviors and psychologically influential interpersonal processes that takes place both within social groups (intragroup dynamics) and between social groups (intergroup dynamics) [4].

Here we focus on the former aspect of group dynamics, and in particular we apply intragroup dynamics approaches to the estimation of the decision-making behavior of small groups through conversation. In prior group dynamics studies, it has been shown that the characteristics of each group vary based on its classification and status, even in relation to decision-making among members [3,4,6,8,9].

### 2.2 Group Decision-Making

Group decision support systems using electronic communication and other computing methods have already been investigated in previous research [2]; however, our study focuses on the actual speech used in group conversations, and how group decision-making can be supported mechanically by the provision of information or suggestions through a conversational agent system.

### 2.3 Utterance Analysis

We define an *utterance* as the smallest unit of speech of spoken language, that is, a continuous piece of speech beginning and ending with a pause, *speech* as the vocal form of human communication, *conversation* as a form of interactive, spontaneous communication between two or more people, typically occurring

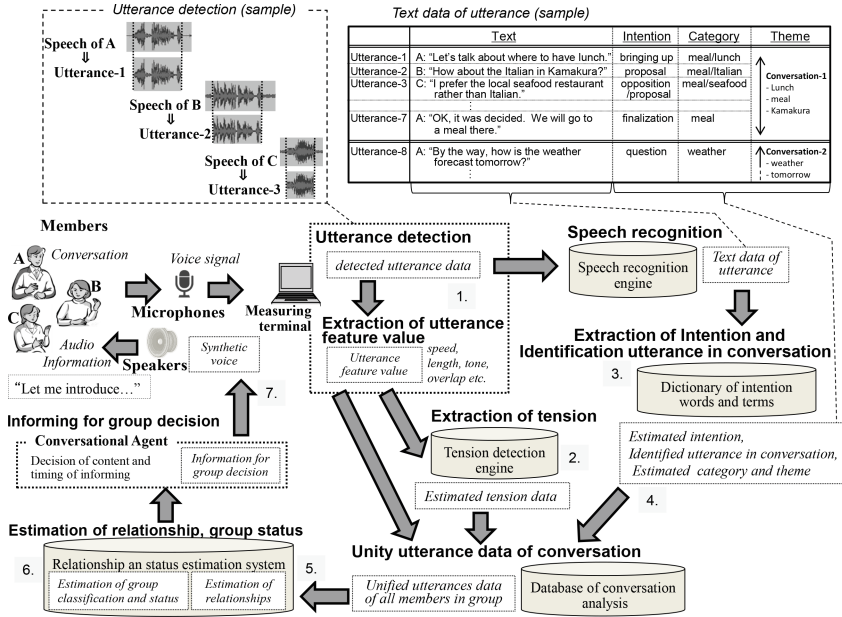


Fig. 1. Processes of group status estimation and provision of group decision-making. \*Numbers 1 to 7 refer to the descriptions provided in the main text.

in spoken communication, and a *conversational agent* as a computer system intended to converse with humans.

Utterance feature values such as the spectrum of utterance power levels have been used for estimation of the status and the tension of members in previous research [1]. *Tension* is defined as mental manifestations of physiological responses in this paper.

We define *intention in spoken dialogue* as a plan or an expectation in a speaker’s mind to do something that has been mentioned in their speech, and it can be estimated by comparing the text data between speech recognition results and spontaneous dialogue corpora. Methods of intention extraction in spoken dialogue utterances have been established by prior research [5].

### 3 Methodology

The methodology we use here, which is based on real-time reactions to speech and its responses, is not new, as noted in related work. However, our proposed method aims to go beyond previous work and recognize the group status to find the most appropriate answers; that is, it does not simply provide answers to speech.

### 3.1 Extraction of Intentions in Utterances

#### 1. Voice signal monitoring and utterance detection

First, each member's voice signals are monitored and differentiated from the voices of other members and background noise. Then, the voice signals are segmented one by one, and the utterance feature values are extracted from the dialogues (Fig. 1-1).

#### 2. Extraction of intentions and tension of utterances in a conversation

Each utterance in speech is recognized, and its content is converted to text. Then, the speed of the utterance (mora/msec) is calculated using the speech recognition results. After the speech recognition process, "subject (theme)" (e.g., the place to eat lunch, the gift to buy in the shopping mall), "category" (e.g., meal, shopping), "sentence style" (e.g., positive, negative, interrogative), "intention" (e.g., proposal, question, agreement, opposition), and "expected action" (e.g., decision on where to have lunch, searching the shop to buy gifts) are extracted by comparing the text data from the speech recognition results and spontaneous dialogue corpora data (Fig. 1-3). Examples of utterance feature values and text data for group status estimation is shown in Fig. 2. The tension levels are also extracted by a calculation using the extracted utterance feature values (Fig. 1-2).

#### 3. Estimation of relationships among members and classification and status of the group

Then, by extracting the content of the utterances between group members in step 2, we can also estimate the relationships among those members (Fig. 1-5). These data also serve as reference information for the estimation of the classification and status of the group (Fig. 1-6). Sample data showing the relationships among group members are shown in Fig. 6.

### 3.2 Decision Support Based on Group Status

We proposed an approach to enhance the overall group condition through the intervention of a conversational agent system using a synthetic voice at an appropriate time based on the group's estimated classification and status as shown in Sect. 3.1. We hypothesize that this intervention will lead to more satisfactory decision-making results. Our proposed intervention examples for the various group classifications are as follows. The agent provides information to a target member through the influencer like a facilitator in conversations, if such an influencer exists. Otherwise, the agent provides the information to a target member directly, if no such influencer exists.

#### 1. High intimacy and flat relationship group

The members of this group classification are assumed to share their opinions frankly. For example, it can be determined whether the members have any specific ideas or requests, and then the conversational agent can provide detailed information based on the situation of each member.



Input	Extracted items	Items for estimation of relationship between members	Items for estimation of group status	Items for estimation of group classification
<b>Voice signal of utterance</b> 	<b>Utterance feature value</b> -tone -speed -power level -length -times -overlaps ↓ <b>Tension</b> -in positive reply -in negative reply etc.	-strength of connection and intimacy -existence of hesitating -hierarchical relationship -parent-child relationship	-existence of leader -existence of elders -existence of subgroup -satisfactions for decision -frequency of speech	-level of intimacy -existence of flat connections -existence of hierarchy
<b>Text data of utterance</b> 	-intension -theme, category -polite word -instruction word -childish word	in each link	in a group	

Fig. 2. Examples of utterance feature values and text data for group status estimation.

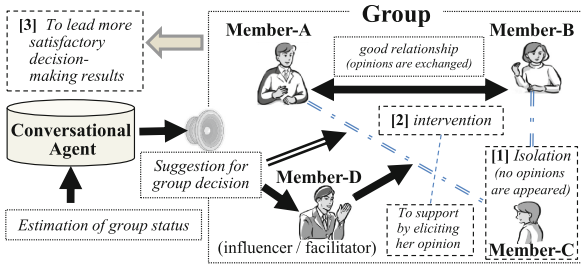


Fig. 3. Group intervention in our proposed model.

2. High intimacy and hierarchical relationship group

This group classification assumes that an older member has the leadership role and knows the views of each member. For example, the dialogue may start with the conversational agent asking the older member what kind of information is preferred by all the members. Then, the conversational agent can provide the appropriate information.

3. Low intimacy and hierarchical relationship group

This group classification assumes that the older member controls the group and junior members may be hesitant to express their feeling directly. It identifies members in a weak position, for example, those who experience isolation from other members, and aims to support such members by eliciting their opinions using appropriate reference information or suggestions (Fig. 3).

We consider applying this logic of supporting the group decision-making process (e.g., destination, venue) while traveling in a car as a test case. We aim to lead the group to higher overall satisfaction by providing appropriate

reference information and suggestions in the discussion prior to a group decision through the conversational agent system using the estimated classification and status of the group.

In our proposal, we focused on supporting members in the weak position mentioned above rather than those in the strong position in the discussion [7]. The procedure uses voice signal monitoring to infer the status of the group and determine the provision of information (Fig. 1-7).

Then, we assume that the connections of utterances and relationships between participants in the conversation can be estimated using the above information and replacing any missing information with utterance characteristics (Fig. 1-5). We prepared a method of visualizing the utterance statuses for the purpose of verification, as shown in Fig. 4.

Then, we inferred the classification and status of the group by measuring the utterance characteristics of members in group, and provided information and suggestions based on an estimation of the group’s status (Fig. 1-6).

### 3.3 Prototype System Configuration

We manufactured a prototype system to verify the operations and methods used in the abovementioned test case in which the group members discuss decisions to be made while traveling in a car. The system configuration of the prototype is shown in Fig. 5. This prototype is composed of three functional parts: (1) utterance measurement, (2) group status and decision estimation, and (3) provision of information for group enhancement.

## 4 Preliminary Experiment

The prototype system was developed, and the basic operation of the system was tested using analysis of utterances and estimations of relationships between members in conversations during preliminary experiment. We have not prepared the appropriate noise reduction system for a car as yet, thus we implemented the test in a conference room, which was hardly affected by noise at all. In the next step, we will prepare the appropriate noise reduction system and implement the test in a car.

We analyzed four kinds of test dialogues with a three-member group (Group A: age 30–55, all males) and a two-member group (Group B: age 25–49, both males) in April 2015. These dialogues each lasted for approximately 5 min, and we processed these test dialogues using the prototype system. The themes of the test dialogues were a sports event and a sight-seeing tour in Tokyo for Group A, and a summer festival and a party for Group B. The results of the preliminary experiment are shown below.

1. Speech recognition of words for extraction of intention  
 Speech recognition was 78 % on average, as shown in Table 1. This is not a particularly high percentage, but it is sufficient to extract the intention of almost all conversations.

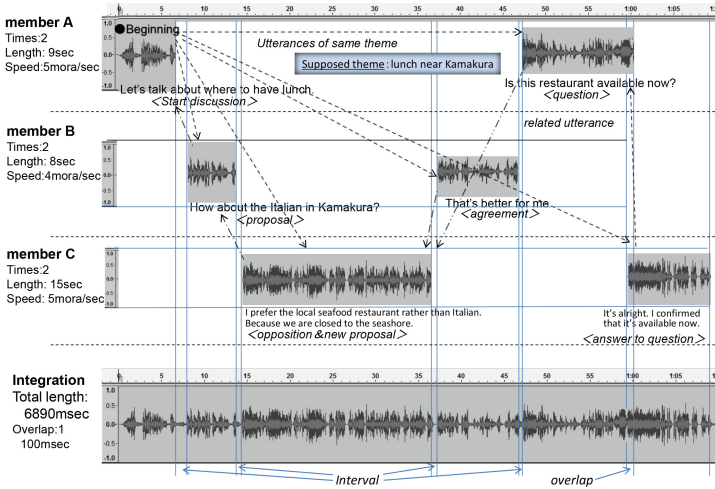


Fig. 4. Image of visualized utterance for verification.

Table 1. Results of the preliminary experiment.

Conversation #	Group name	Numbers of utterances	Recognized utterances including target words (value, rate)	Extraction of utterances (value, rate)	Extraction of utterances in conversation (value, rate)
1	A	74	19/29 65.5 %	19/69 27.5 %	69/69 100.0 %
2	A	70	20/25 80.0 %	20/65 30.88 %	65/65 100.0 %
3	B	47	18/21 85.7 %	18/45 40.0 %	45/45 100.0 %
4	B	65	21/25 84.0 %	21/61 34.4 %	61/61 100.0 %
Average	—	64	78.0 %	32.5 %	100.0 %

2. Extraction of intention of utterance

The intention of utterance could be extracted in only 32% of all utterances as shown in Table 1. However, we could identify all utterances in conversation through the extraction of intention and time course information. It is necessary to increase the number of words in the dictionary for more appropriate extraction of intention.

3. Extraction of the utterance feature values

It was quite successful, and we could calculate the strength of the tension using these utterance feature values. We confirmed that the strength of the tension was correctly estimated through human monitoring, excluding any utterances that did not have sufficient length or power to judge the tension.

4. Estimation of group classification and status

The status of utterances between members in a test conversation is shown in Table 2, and the connections between members is shown in Fig. 6. The group classification and status could also be estimated using these estimated

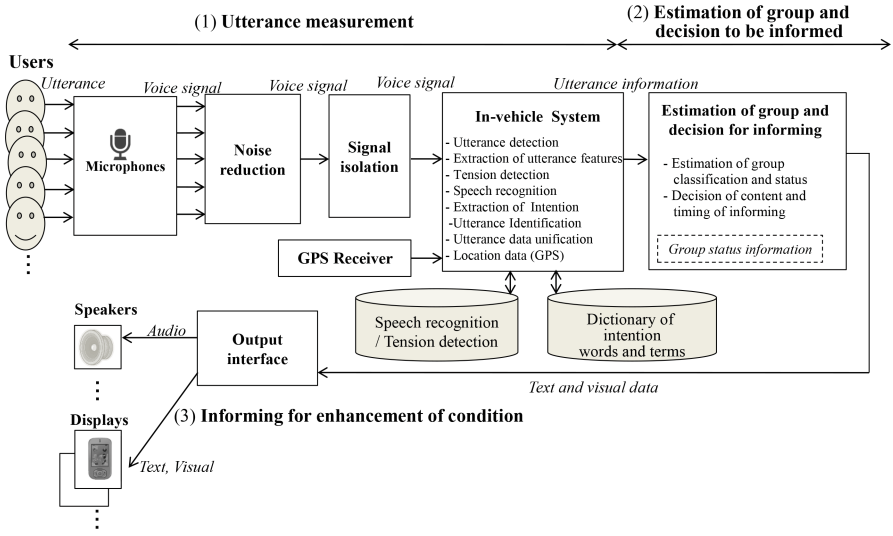


Fig. 5. System configuration of prototype.

Table 2. Status of utterances between members in a test conversation.

Link no.	Utterance direction	Numbers of utterances	(Details of intentions)
1, 5, 7, 9	A → B, C, D, E	14	(question:10, proposal:1 etc.)
2, 3, 4, 8	B → A, C, D, E	4	(proposal:4)
2	B → A	1	(agreement:1)
10	C → A	5	(negation:3, agreement:1 etc.)
6	D → A	2	(agreement:1 etc.)
11	E → A	8	(agreement:2)

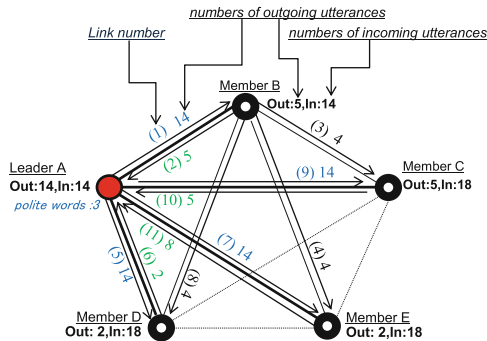


Fig. 6. Connections between members in a test conversation.



relationships and connections and the collected data (e.g., utterance feature values, intention, strength of tension).

While the degree of agreement between the estimated group status and the real group status remains insufficient, we plan to further optimize this procedure by testing using many additional kinds of utterance data in future experiments.

5. Judgment of contents and timing of provision of information

We did not implement the function of informing members in this preliminary experiment. However, we confirmed that extraction of utterance data for judgment of contents and timing of provision of information is possible. We plan to further optimize this method by continued testing with many additional kinds of utterance data in future work.

6. Degree of Satisfaction with Decision-Making Conversation

We aimed to measure the index of satisfaction in decision-making conversation by understanding how members express their own opinions, how those opinions are discussed with other members, and how the final group decision is reached. We obtained this index of satisfaction by extracting the intention of utterances at the current stage. For example, the intention was calculated using the number of times keywords were expressed in an individual's opinions. However, merely extracting keywords from each utterance is not enough for the accurate estimation of satisfaction, thus we will continue to study the proposed method.

## 5 Conclusion

In this paper, we proposed a system to infer the classification and status of groups by measuring the utterance characteristics of the group members, and to enhance the overall group condition through a conversational agent system based on estimations of group dynamics.

The basic availability of our proposed method to measure the utterance characteristics of group members was confirmed by our prototype system in our preliminary test. In future work, we will further optimize the logistical and system functions for estimating group status and the continuous provision of information to the group members.

We will also proceed to further verify the details of our method and system through continuous field tests, and we hope to verify the ability to increase the group members' satisfaction with the group's decision-making process through the use of our proposed conversational agent. Specifically, in our future work, we will collect many additional kinds of utterance-test data and further clarify the appropriate parameters for estimation and information provision using machine learning.

## References

1. Ariga, M., Yano, Y., Doki, S., Okuma, S.: Mental tension detection in the speech based on physiological monitoring. In: IEEE International Conference on Systems, Man and Cybernetics, ISIC, pp. 2022–2027. IEEE (2007)

2. Desanctis, G., Gallupe, R.B.: A foundation for the study of group decision support systems. *Manag. Sci.* **33**(5), 589–609 (1987)
3. Deutsch, M., Gerard, H.B.: A study of normative and informational social influences upon individual judgment. *J. Abnorm. Soc. Psychol.* **51**(3), 629 (1955)
4. Forsyth, D.: *Group dynamics*. Cengage Learning (2009)
5. Hodjat, B., Amamiya, M.: Applying the adaptive agent oriented software architecture to the parsing of context sensitive grammars. *IEICE Trans. Inf. Syst.* **83**(5), 1142–1152 (2000)
6. Hogg, M.A., Tindale, S.: *Blackwell Handbook of Social Psychology: Group Processes*. Wiley, New York (2008)
7. Kono, S., Aihara, K.: A model of decision support based on estimation of group status by using conversation analysis. In: Stephanidis, C., Tino, A. (eds.) *HCI 2015 Posters*. CCIS, vol. 528, pp. 627–632. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-21380-4\\_106](https://doi.org/10.1007/978-3-319-21380-4_106)
8. Kelley, H., Thibaut, J.: *Interpersonal Relations: A Theory of Interdependence*. Wiley, New York (1978)
9. Lickel, B., Hamilton, D.L., Wierzchowska, G., Lewis, A., Sherman, S.J., Uhles, A.N.: Varieties of groups and the perception of group entitativity. *J. Pers. Soc. Psychol.* **78**(2), 223 (2000)