Global Interactional Patterns in Group Oral Interaction

Junko NEGISHI
Waseda University

Abstract

This study aimed to investigate the discourse of 135 Japanese learners of English ranging from junior high school to university students, in terms of their interactional patterns while interacting in a group of three. The participants’ five-minute interactions were assessed by ten Japanese raters by means of the rating criteria in Common Europe Framework of Reference (CEFR; Council of Europe, 2001). The relationship between the interactional patterns and the raters’ assessment was also explored. Four “Global Interactional Patterns” including a newly observed pattern termed under-developed, were confirmed in addition to the three basic patterns of interaction that Galaczi (2004, 2008) advocated: collaborative, asymmetrical, and parallel. The difference in the results was likely to have stemmed from the dissimilarity in participants’ speaking ability. As for the relationship between the raters’ assessment and the Global Interactional Patterns, the collaborative pattern yielded the highest scores and under-developed the lowest. When the author assigned scores to both group and individual characteristics (e.g., three points to collaborative interaction and one point to a passive speaker), a strong correlation (.607, p < .01) was observed but the correlation was largely explained by group rather than individual characteristics.

1. Introduction

1.1 Interaction

Kramsch (1986) introduced the concept of “interactional competence” which originally stemmed from a sociolinguistic point of view. The concept regards collaborative interaction as important because it subsequently prompts second language (L2) acquisition. He and Young (1998) and Young (2000), who took up the notion as Interactional Competence Theory, view interactional competence as co-constructed by all interlocutors in a discursive practice.

1.2 Oral performance tests: interviews and group oral interactions

Interviews have been commonly used for oral performance tests; nevertheless, this test format is often criticized as being “pseudo-social and asymmetrical” (van Lier, 1989). In contrast to the interview test format, paired or group oral interactions (“group oral interaction” will be termed “group oral” hereafter, following Bonk & Ockey [2003]) have rarely been implemented as a performance-based test since they have been believed to have various defects. The first
drawback concerns the quality of the interlocutor who is not a trained or professional interviewer but is a test-taker and may disrupt the other interlocutors in the group (Brooks 2009). Consequently the interaction might be affected by uncontrollable variables from other peers which threatens its validity. Second, compared with an interview, an assessment of a paired/group oral interaction generally demonstrates lower inter-rater agreement, which is also a threat to validity (Van Moere, 2006). In addition, there is still controversy over whether raters should give scores to each member of the pair/group or to the group as a whole. Representing the interactional competence point of view, May (2009) argues that raters may have difficulty in reducing the influence of one speaker on another in order to provide separate and fair scores to individuals. Meanwhile, many researchers (cf. Bonk & Ockey, 2003; Van Moere, 2006) think it appropriate to test in groups but to rate test-takers as individuals. These disadvantages, which may pose a threat to test validity, might have been the reason why the test formats have been used rarely.

Though there are various deficits, the group oral also has advantages. First, the group oral interactions promote more communicative, speaking-focused learning and teaching, which may reflect classroom situations and induce the washback effect, because collaborating during speaking activities helps develop the learners' oral proficiency. Second, oral interaction enables speakers to show more varied interaction patterns and language functions within a richer discourse, which leads to authenticity (cf. He & Dai, 2006). In this regard, speakers have more opportunities to show their oral ability or interactional competence as they employ more interactional and conversational management functions (Taylor, 2001; Fulcher, 1996; Galaczi, 2004). The more authentic the test task becomes, the more valid the score interpretation (Ockey, 2009). Third, oral interaction evens out the asymmetrical power relationship observed in the interview test format (cf. Taylor, 2001). Fourth, candidates are reported to be able to earn higher scores in a paired test than in an interview (Brooks, 2009). Bonk and Ockey (2003) state that “providing the students with opportunities to initiate and control conversation during the test might mean an enhancement of the validity of the score-based inferences” (p. 90). Fifth, test-takers show positive reactions to the group oral; they consider such discussions with other peers less intimidating than a face-to-face interview, in that they can control the conversation and are allowed to use more natural language (cf. Van Moere, 2006). Sixth, raters can concentrate on their assessment; they need not participate in the interaction with the examinees eliciting test-takers' speech sample. This also means that rater training is less difficult than training an interviewer. Last, the group oral is resource economical (Bonk & Ockey, 2003). In particular, the format enables educational institutions to test many students at a single time. In fact, the group oral test format has been implemented for university entrance exams in South Korea for its economic efficiency and for time-saving reasons. Accordingly, Korean students practice how to interact in English, and believe it promotes pragmalinguistic competence (personal opinion by a committee member).

For these reasons, dyadic or group oral test formats have recently been more employed: for example, in Cambridge First Certificate (paired), Cambridge Certificate of Proficiency in English.
(paired) and in the speaking test administered by the Council of Europe (paired and group oral). There are some local tests that utilize the group oral. In South Korea these are run by universities (for entrance examinations as mentioned earlier), the Educational Testing Service (ETS; to select scholarship recipients) and a private company; in Hong Kong there is the Hong Kong Use of English test, and in China there are the College English Test–Spoken English Test (CET-SET) and the Public English Test Systems. No group oral tests have been administered in Japan to date, excluding an English Forum for junior and senior high schools organized municipally, whose purpose is to encourage and improve students’ communicative competence as advocated in the Japanese course of study.

1.3 Research on paired and group oral

Compared to research in interview tests, studies on paired and group oral interactions have been very limited. Some research concerning the paired oral, which can be partially applied to the group oral, has reported that 1) just as interlocutor effects can be found between an interviewer and interviewee, so can they be found between interlocutors in a paired format (Brooks, 2009), 2) interlocutor effects seem to be indirect and unpredictable (Brown & McNamara, 2004), and 3) demonstration of collaborative interaction might have indicated higher oral interaction skills (Glaczi, 2004; 2008).

Though limited, there is some research concerning the group oral. Ockey (2009) investigated interlocutor effects in groups of three, studying whether scores assigned to assertive and non-assertive speakers were affected by interlocutors’ levels of assertiveness. The assertive speakers received higher scores when grouped with the two non-assertive interlocutors and lower scores when grouped with the two assertive interlocutors. He and Dai (2006) investigated 60 groups of Chinese learners of English in the CET-SET group discussion administered by the National College English Testing Committee. The test was carried out face-to-face by three group members. According to the researchers’ observation, the candidates tended to talk to raters rather than to interlocutors. Accordingly, He and Dai reported that low interactional language functions (ILFs), such as developing and negotiating meaning, were observed. It is likely that the candidates put too much emphasis on the speaking test and, consequently, were not aware of the indispensability of meaningful interactions.

The group oral has been reported to threaten the reliability of test scores; however, further research should be carried out because the research is insufficient in this area, compared to the research in interviews. Van Moere (2006) states that “the potential advantages to group assessments, and the fact that candidates themselves give positive reactions to this test format, suggest that further research is warranted into group oral tests” (p. 436). In 2009, one journal, Language Testing, featured various researches concerning paired (and group) orals, which indicates the beginning of a new era of the test formats.
1.4 CEFR rating scales

Council of Europe (2001) has released its Common European Framework of Reference for Languages: learning, teaching, assessment (CEFR hereafter). The Manual of the CEFR (Council of Europe, 2003) gives two kinds of rating scales for the paired/group oral: CEF global oral assessment scale (p. 78) and CEF oral assessment criteria grid (p. 79). What they have in common is that there are 6-level “Can Do” descriptors, that is, A1, A2 for Basic Users, B1, B2 for Independent Users, and C1, C2 for Proficient Users. The difference between the two rating scales is that the former sees an examinee’s performance from a holistic point of view and the latter has 5 subcategorized analytic grids: Range, Accuracy, Fluency, Interaction, and Coherence.

1.5 Purpose of the study

Speaking assessments in Japan are meant to facilitate the use of picture description tasks or interviews by individual examinees, while group oral discussions are rarely carried out. Given this factor, this study aimed first to find some features and developmental phenomena of the 135 English learners’ discourse in terms of their interactional patterns when interacting in a group of three. Second, the relationship between the rating scores given by raters based on the CEFR rating criteria and the participants’ interactional patterns was investigated. This paper solely reports the results of investigation in Interaction among the five rating subcategories on the CEFR rating criteria.

2. Method

2.1 Participants and settings

The participants in the study were 135 students from seven schools from among three kinds of educational institutions, namely, two junior high schools, two senior high schools, and three universities in and around Tokyo. They were divided into a total of 45 groups, each containing three students. The groups comprised fifteen junior high school student groups, fifteen senior high school student groups, and fifteen university student groups. In order to ensure an appropriate balance of students from the various types of schools, about half of the participants were recruited from public schools, while the others were from private schools. The university students belonged to a wide-range of faculties, none of them being English majors. No students had received education abroad with English as the medium of instruction.

The data on the group oral were collected from each educational institution through the following process: 1) The students were randomly allocated into groups of three; 2) Each group drew a card on which one of the six interaction topics—School, Family, Friends, Hobbies, English, and Culture (the last being only for university students)—was written down, and they were asked to speak on the topic; 3) Five minutes were allotted to each member of the group to plan his/her speech without speaking to the other members of the group; 4) Each member of the group
introduced themselves for about half a minute as a warm-up activity; 5) Finally, the three students interacted orally as a group for five minutes on the selected topic. They were encouraged to have a natural and casual conversation while sitting and looking at each other. The interaction was videotaped after acquiring the permission of the participants. All conversations were transcribed with the aid of DVDs which compensated for the deficiency of the information.

2.2 Ratings and multi-faceted Rasch analysis

The participants’ oral interaction was assessed by ten Japanese teachers of English who had master’s degrees in the field of English education or linguistics. Before rating, they were trained using a training DVD accompanied with the CEFR Manual (Council of Europe, 2003). They rated the students while watching the participants’ DVDs by applying both a holistic rating scale and analytic rating criteria by seven scales: Below A1, A1, A2, B1, B2, C1, and C2.

In most oral performance tests for second language learners, raters assess participants’ speaking ability by means of rating scales. No matter how precise the rating scales are, the assessment might be subjective or inconsistent because the ratings are given by humans (McNamara, 1996). Training is mandatory to reduce irregular errors and improve inter-rater reliability and intra-rater consistency; however, it is impossible to remove all excessive severity or leniency of the ratings (McNamara). Multi-faceted Rasch analysis compensates as much as possible for diverse influences, which have many causes or “facets” such as raters’ differences, task difficulty, examinees’ abilities, and their interactions. It estimates and transforms original “raw scores” given by raters to “measures,” which are on an interval scale. This study utilized multi-faceted Rasch analysis software, FACETS, Version No. 3.6.4.0 (Linacre, 2008) to eliminate subjective judgment as far as possible.

2.3 Analysis of interaction

In a paired or group test format that requires interlocutors, as variability increases, the more difficult it is to establish the test construct, reliability, and fairness. Such conditions have brought about the shift from a cognitive perspective to a social perspective of interaction (Brooks 2009). The use of qualitative analysis has been increasing, and more researchers are utilizing qualitative methods to complement their quantitative analyses for pursuing validation verification. Among the researchers, Storch (2002) and Galaczi (2004, 2008) applied the methods of conversation analysis (CA) in order to explore L2 learners’ discourse in a paired test.

Storch (2002), for the purpose of investigating speakers’ patterns of interaction in an adult ESL classroom, has constructed a model of dyadic interaction, shown in Figure 1. The model introduces four role relationship patterns: collaborative, dominant/dominant, dominant/passive, and expert/novice, based on two concepts, equality and mutuality. According to the model, mutuality refers to “the level of engagement with each other’s contribution,” and equality refers to “the degree of control or authority over the task” (Storch, 2002, p. 127).
Galaczi (2004, 2008) has adapted the Storch’s dyadic model (2002) for the purpose of analyzing the discourse produced by the test-takers of Cambridge First Certificate of English speaking test (FCE). Figure 2 is the author’s schematic representation of Galaczi’s patterns of interaction, which are described in detail in three tables. There are four interactional patterns: collaborative, parallel, asymmetrical, and blended. According to Galaczi’s analysis, collaborative interaction resulted in the highest Interactive Communication (IC) scores in the FCE, whereas parallel interaction resulted in the lowest. In this study, the participants’ discourse was analyzed based on Galaczi’s three patterns which excluded blended patterns.

3. Results and Discussion

3.1 Scores and “measures” obtained from the multi-faceted Rasch analysis

The scores given to the participants by the raters using the CEFR rating criteria were put into the multi-faceted software, FACETS. Although FACETS analysis provided varied information such as task difficulty, participants’ ability, rater severity and so forth, this study solely used the produced interval scores or “measures” expressed in “logits” for one of the subcategories, Interaction (For details, see Negishi [2010] in ARELE). An average measure of junior high participants was -6.27 logits (approximately A1 for CEFR assessment), senior high was -2.40 logits (A2), and university was 0.06 logits (B1), all of which showed a clear difference of the participants in terms of which educational institution they belonged to.
3.2 Global Interactional Patterns

To explore the participants' discourse, the analysis was carried out based on Galaczi's (2004, 2008) classification described in section 2.3; however, this study was different from that of Galaczi, whose participants were upper-intermediate candidates taking the FCE test therefore able to develop the assigned topic more easily. The students in this study, on the other hand, ranged from novice to intermediate and needed to elaborate on the topic solely to maintain their interaction. Given this situation, the patterns of interaction that Galaczi advocates above did not fully apply to the current study. The novice speakers in this study had difficulties in carrying out or developing the assigned topics coherently over several turns and consequently often changed subordinate topics abruptly without any cues. This newly observed interactional pattern was termed *under-developed* interaction, shown as No. 4 in Figure 3. For this reason, the model in this study includes *development* in place of *equality*. *Development* in this context refers to mutual topic development among the interlocutors. *Equality* did not necessarily play an important role among the participants because they belonged to the same educational institutions and they were at similar levels. The participants categorized as under-developed interacted mutually, trying to elicit other speakers' utterances and requesting more information without success, owing to the fact that they did not know how to respond or extend contributions by the prior speakers. Consequently, these speakers were compelled to change the topic out of confusion, which is less common among advanced speakers. The under-developed speakers frequently used their native language, Japanese, and laughed when they were at a loss. Another feature of their talk was that each turn and utterance was short, and the utterance often lacked subjects or objects.

![High mutuality diagram](image)

*Figure 3. Global Interactional Patterns, including under-developed interaction*

*Note: Numbers in parentheses refer to the proportion of the patterns and the average CEFR measures given to the participants.*
Other three interaction patterns that Galaczi (2004, 2008) advocated were also observed in this study. Parallel interaction showed low mutuality and less developed style in terms of expanding the assigned topic among the interlocutors. A speaker may develop a self-initiated topic on his/her own, but not mutually or collaboratively. Asymmetrical interaction also had low mutuality, but the most assertive speaker developed the topic in a less collaborative manner. Collaborative interaction exhibited a mutual exchange among the group members in order to develop the assigned topic, which was the ideal type of interaction. These four patterns were termed “Global Interactional Patterns.”

The participants’ discourse was classified as one of the four patterns shown in Figure 3, based on the dominant pattern so that there were no blended patterns, in contrast to Galaczi’s study (2004, 2008). Blended patterns may complicate the classification, because this study is about the group oral, rather than the paired. Overall, 37.8% of the interactions were categorized as under-developed, 20.0% as parallel, 6.7% as asymmetric, and 35.6% as collaborative, as shown in Figure 3. Galaczi’s classification showed 30% collaborative, 30% parallel, 10% asymmetrical, 23% collaborative-parallel blended, and 7% collaborative-asymmetric blended. The results of the two studies are not comparable because the current study does not include the blended patterns. However, both this study and that of Galaczi found that the groups assigned the highest scores were collaborative and those with the second-highest scores were asymmetric, while the parallel received the lowest. The under-developed interactional pattern resulted in much lower scores than the parallel.

Figure 4 depicts the Global Interactional Patterns sorted by the educational institutions. With regard to the junior high school groups, most of the interactions are categorized as under-developed, excluding two collaborative interactions. In the public senior high school groups, mixed interactional patterns are observed: three under-developed, one parallel, and four collaborative. Interestingly, in contrast to the public senior high school groups, the private senior high school groups have only a parallel interactional pattern. As for universities, the public university groups have two collaborative and three asymmetric patterns.
which are seen only among the public university groups. One of the two private universities has varied patterns: one under-developed, one parallel, and three collaborative. Another private university, the most proficient of the three universities, exhibits only collaborative patterns. As regards collaborative patterns, the junior high school groups have only two, the senior high groups have four, and the university groups have ten out of fifteen, indicating that the more proficient the speakers become, the more they tend to be collaborative.

The following excerpts are examples of the four types of Global Interactional Patterns.

Excerpt 1 (JHS: group 11)

9  L: Because I like tennis. Is the club fun?
10 →M: Yes, yes, it is. How about you.
11  R: Yes, it is.
12  L: (inaudible)
13  M: Do you play tennis well.
14 →L: No, I don’t. How about you.
15 →M: Mmm, no, I ain’t, don’t. How about you.
16  R: (inaudible) No, I don’t.
17  M: <<Japanese words>>
18  L: Do you like your club, T***** <<L’s name>>.
19 →R: Yes, I, yes, I do. How about you, M***** <<R’s name>>.
20  M: Mmm, I, mmm, I’m... Yes, I do.

Excerpt 1 is a typical example of the under-developed interaction carried out among one of the public junior high school groups. They seem to avoid questions that require them to deepen their thought. Instead, they keep asking simple questions using “How about you?” (indicated by → and underlines). Accordingly, in this short interaction, they take turns swiftly and frequently. These quick turns are likely to put less of a burden on each student. This type of recycling of questions is common among the junior high groups. Most of the under-developed interactions include such identical questioning styles, demonstrating their willingness to yield the floor. Another common feature of the under-developed interactions is the high mutuality of the speakers.

Parallel interaction was carried out with low mutuality and a less developed manner while the speakers were trying to expand the assigned topic. The participants who engaged in parallel interaction extended their self-initiated topics without being mutually collaborative. As shown in Figure 4, the private senior high school is the single educational institution that displays the largest number of parallel interactions. The reason for this can be hypothesized as follows: the learning style of the private senior high school is slightly different from normal senior high schools; the private senior high school puts a greater emphasis on grammar and writing than on speaking, for the purpose of preparing students for university entrance examinations. What is required of
students is that they be capable of understanding or writing English, not necessarily communicating with other people. It is likely that they express themselves in English as if they were writing in solitude. The following is an example of the parallel interaction in the private senior high school groups:

Excerpt 2 (SHS: group 25)

1 M: Uh, what is your hobby, K***** <<R’s name>>.
2 R: My hobby is reading book and reading comic and, and playing TV game and I’m very indoor man so, I love in, I love playing in my house. What your hobby.
3 M: My hobby is playing sports. Uh, especially, especially, I, I like playing basketball or American football. (laughter) So, uh, play, uh, playing basketball is uh, very much eh? Eh: but I, I don’t like read a book, uh uh:: so, so much. How about you, I******* <<L’s name>>.
4 L: My hobby is to playing tennis, play tennis. Uh, eh-toh, <<Japanese words>> Mmm, eh, when I was junior high school student, I was, I belonged to the tennis club. Eh mah, it was fun (laughter) very much.

In Excerpt 2, triggered by M’s question, “Uh, what is your hobby, K*****”, the group members tend to talk only about their own hobbies. R’s hobby is reading, while the second interlocutor, M, mentions sports. We can find the single relevant comment when M refers to the prior speaker’s utterance: “I, I don’t like read a book [sic], uh uh:: so, so much.” The third speaker does not refer to the utterances of the other two interlocutors. Their conversation continues in this style.

All three asymmetrical interactions were found only in the public university. Not only the group whose interaction was excerpted below but also the other two groups had one talkative speaker, supported by the other two interlocutors. Despite the talkativeness of the dominant speakers, they were not considered proficient, because their utterances contained various mistakes, hesitations, and repetitions. It was a phenomenon that the three groups had in common and that gave rise to unintelligibility.

Excerpt 3 (U: group 35)

7 R: Aah huu huh huh, yeah. And so, so, l***** and T**** <<prefecture names>> is uh, un, same, same words intonation. And, en, mmm, cou-, in a count, country words as similar to l***** and T***** and F******* <<another prefecture name>>. Do you know?
8 ➔M: Sorry? Pardon?
9 R: Hooh hoo hoo? Uh, so l***** and T*****’s, eh, countrywords is similar. Uh, so, mmm, for example, and, uh, so mid, so uh, deresuke and gojyappe and aonajimi, do you know? Do you know Y***** <<L’s Name>>?
10 L: Uh, yes=
11 R: =Yes.
12 ➔L: I think aonajimi is, uh, uh, Tokyo word is aozaa=
13 R: =Yes, yes, so yes.
14 ➔M: It’s very painful.
Excerpt 3, R’s utterance in Turn 7 is unintelligible, and M has to ask for clarification by saying, “Sorry? Pardon?” in Turn 8. Although M asks for clarification, R does not seem to feel sorry for being incomprehensible; rather, he regards M as a poor listener. The other two interlocutors, L and M, try to acknowledge, develop, support, and challenge R’s talk (indicated by \[\rightarrow\] and underlines); however, the lack of balance is obvious. R initiates and extends his topic, and the others do not attempt to take long turns, which may have caused the asymmetrical interaction.

As mentioned earlier, very few junior high school groups were categorized as collaborative, which indicates difficulties developing the assigned topic collaboratively. Among the senior high school groups, about one-fourths of the groups were categorized as collaborative. Although their English was still poor, they attempted to collaboratively support and draw out other members’ voices. Two-thirds of the university groups were classified as collaborative, the largest number of all educational institutions. Specifically, all the groups from one of the private universities were labeled as collaborative.

Excerpt 4 (U: group 44)

10  M: So um, I think Chinese university students study more. And American, and American students, um, pay, um, pay, uh [mm?]
11  \[\rightarrow\] R: [Fee?]
12  M: No, mmm, pay money. For university by, by, mmm=
13  \[\rightarrow\] R: =By themselves?
14  M: Yes, yes.
15  R: Mmm, that is the mmm, very difficult, uh umm, very different [from Japanese students?]
16  L: MMM. Mmm yeah. I agree?=
17  M: =Mm mmm?
18  L: But now it’s getting better than before. That now the students study much? Harder than uh: they did like forty years ago or something because getting job is very difficult [these days so...]
19  M: [Mmm
20  L: [It’s good.
21  \[\rightarrow\] R: [Because of depression?=
22  L: =Mmm. Maybe.
23  M: Mmmm. Woman students uh like to study a special topic.
24  L: Special topic.
The last excerpt, Excerpt 4, is an example of the university students’ collaborative interaction. Fundamental features of the interaction resemble those of the junior and senior high school groups; for example, the quantity of utterances among the three participants is well-balanced, and follow-up questions or acknowledgements are frequently used (e.g., “Yes, yes.” “Mmm...” “It’s good” in Turn 14, 19, and 20). Overlaps (indicated by [ ) and latches (indicated by =) are repeatedly used, exhibiting the interlocutors’ support and involvement. The students remain on the topic over many turns, while developing the current subordinate topic and smoothly shifting to another subordinate topic. What makes the interaction collaborative is the collaborative floors (Coates, 1996), in which other interlocutors try to complete the prior speaker’s utterances (indicated by → and underlines). This collaborative floor among the group members may be an emphatic exemplification of their involvement and support.

3.3 Relationship between CEFR measures and Global Interactional Patterns

Two types of relationship between the CEFR measures and Global Interactional Patterns were explored. As displayed in Figure 3, the average CEFR measures which the Global Interactional Patterns obtained demonstrate a clear difference: Collaborative yielded the highest measures (−0.39 logits), asymmetrical the second highest (−0.55 logits), parallel the second lowest (−1.65 logits), and under-developed the lowest (−5.37 logits). The discrepancies in the CEFR measures between each pattern on average were 0.16 logits between collaborative and asymmetrical, 1.10 logits between asymmetrical and parallel, and 3.72 logits between parallel and under-developed, last of which showing the largest disjuncture.

<table>
<thead>
<tr>
<th>Table 1 Types of Interaction and Assigned Scores and Correlation Coefficients (Kendall’s Tau)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Global Interactional Patterns</strong></td>
</tr>
<tr>
<td>Assigned Scores</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Correlation between CEFR measures</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

In order to identify whether group characteristics (the Global Interactional Patterns) or individual characteristics (e.g., dominant or passive) play an important role toward the CEFR measures, each characteristic was replaced by a number. For the Global Interactional Patterns, the
same score was assigned to all three members of the group, that is, 3 points for collaborative, 2 for parallel and asymmetric, 1 for under-developed, and no points for numerous breakdowns. As for the individual characteristics, the numbers each speaker allotted were as follow: dominant speakers, 3 points; equal-type and relatively passive speakers, 2 points; and passive speakers, 1 point. These numbers were determined based on the measures the participants received from the raters; for example, dominant students were commonly awarded higher scores than equal-type or passive speakers. Each speaker’s scores for group characteristics and for individual characteristics were added and compared with the CEFR measures. Different score combinations were tried, and the order shown in Table 1 proved to most effectively represent the overall interactional characteristics. However, it should be noted that the replacement has limitations and is utilized solely for descriptive purposes. The rank order correlation coefficient (Kendall’s tau) between the scores of Global Interactional Patterns and the CEFR measures calculated for the group was .586, $p < .01$ whereas the correlation between the individual characteristics and the CEFR measures calculated individually was .225, $p < .01$. This means that the CEFR Interaction scores were associated with the group characteristics rather than the individual characteristics. When adding the scores of Global Interactional Patterns and individual characteristics and comparing them with the CEFR measures, the correlation coefficient between the CEFR measures was .607, $p < .01$. Considering the high correlation between the scores of Global Interactional Patterns and the CEFR Interaction measures, most of the correlation may be explained by the Global Interactional Patterns given to each group, rather than by the individual characteristics.

4. Conclusion

This study explored the characteristics of the participants’ interaction qualitatively and quantitatively in terms of the Global Interactional Patterns. Four types of patterns were confirmed, including a newly observed pattern termed under-developed, which was not in Galaczi’s (2004, 2008) research. The difference in the interactional patterns was likely to be rooted not in the test format, that is, dyadic or triadic, but rather in the dissimilarity of the speakers’ language/speaking ability. Galaczi’s study and this study exhibited similar results in terms of the relationship between the interactional patterns and the scores the speakers of both studies obtained, that is, the collaborative pattern yielded the highest scores, the asymmetrical the second highest, and the parallel the lowest. In addition, the under-developed pattern demonstrated lower scores than the parallel.

The strong correlation coefficient was observed between the Global Interactional Patterns (encompassing group and individual characteristics) and the CEFR measures, .607, out of which the correlation with the group characteristics was .586. Although there is a limit to the effectiveness of replacing the patterns with numbers, this fact suggests the possibility of assigning scores to a whole group, not to an individual. Brooks (2009), in her paired interaction
research, found that the participants’ performance scores were closer to each other when paired than when they performed individually. Brooks thinks that some raters may have awarded joint scores to the paired test takers. The scores that the paired candidates received were better than those for individual performance carried out by the same candidates. The raters of this study showed variability in regard to the dominant/passive speakers, namely, asymmetric interactional participants. This is similar to the problem that May (2009) proposes in her paired speaking test study, in that raters have difficulty in reducing the impact of one speaker on another to assign separate and fair scores. In order to solve this problem, May suggests that joint scores be given for acknowledging the necessity of co-constructed interaction, though she recognizes the difficulty of this practice in the case of high-stakes tests. Considering the fact that the CEFR measures have a significant correlation with the group characteristics in terms of Global Interactional Patterns, the possibility of assigning scores as a group should be considered, specifically in the case of multiple task test formats or the same educational institutions. The research concerning the joint scores will be beneficial for future research.

Despite the fact that there are disadvantages, employing the group oral may be beneficial in developing diverse types of performance test. Bonk and Ockey (2003) state, “The one-shot test of discussion ability is certainly an insufficient basis upon which to make a valid overall decision about an examinee’s oral ability, but it may more closely match the construct of oral L2 proficiency… In many institutions worldwide, performance assessment is still a new concept; … thus, the group oral, along with its limitations, may be a step in the right direction” (p. 103). Fulcher (1996) claims that group discussion is easier for L2 learners than other types of performance tests and may be better suited for less proficient learners. The participants in this study also gave positive feedback to this format. The group oral interaction can be carried out with novice to intermediate participants by providing familiar topics, as this study indicates. This type of interaction can be utilized by educational institutions for the purpose of placement tests or term- or year-end evaluations.

With regard to pedagogical implications, what we need to focus on in oral interaction is that students must know how to interact and to co-construct the conversation between interlocutors. These abilities may be the most important factor in a real-world communication. Such interactional competence cannot be improved solely by monologues, story-telling or picture descriptions. Classroom group oral activity will bring about close relations with the group oral test format with practical communication. In this sense, teachers are being required to link classroom activities and oral interaction tasks in educational contexts. Moreover, the pedagogical focus should not be on grammar alone. Students of the private senior high school, where grammar, writing and reading were mainly taught, could produce longer sentences with fewer mistakes, but they were not as collaborative as the students of other institutions. The analysis of Global Interactional Patterns showed that groups with collaborative interaction received higher scores; in contrast, parallel and asymmetric interaction obtained lower scores. It is thus assured that
collaborative interaction as a classroom activity helps develop learners’ oral interactional competence.

References


Appendix

**Transcript notations**

- Overlapping utterances
- Latching that indicates no interval between adjacent utterances
- Falling intonation: e.g., sentence final
- Rising intonation (does not mean a question)
- Stressed syllable
- A prolonged stretch
- Japanese words
- Inaudible utterance
- Author’s description