



Conversation Proficiency Assessment: A Comparative Study of Two-Party Peer
Interaction and Interview Interaction Implemented with Thai EFL Learners

Ratchawan Ussama

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of
Master of Arts in Teaching English as an International Language

Prince of Songkla University

2013

Copyright of Prince of Songkla University

Thesis Title Conversation Proficiency Assessment: A Comparative Study of
Two-Party Peer Interaction and Interview Interaction
Implemented with Thai EFL learners

Author Ms. Ratchawan Ussama

Major Program Teaching English as an International Language

Major Advisor:

.....
(Dr. Kemtong Sinwongsuwat)

Examining Committee:

.....Chairperson
(Asst. Prof. Dr.Premin Karavi)

.....
(Dr. Compol Swangboonsatic)

.....
(Dr. Kemtong Sinwongsuwat)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Arts Degree in Teaching English as an International Language.

.....
(Assoc. Prof. Dr. Teerapol Srichana)
Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

.....Signature
(Dr. Kemtong Sinwongsuwat)
Advisor

.....Signature
(Ms. Ratchawan Ussama)
Candidate

I hereby certify that this work has not been accepted in substance for any degree, and is not being currently submitted in candidature for any degree.

.....Signature

(Ms. Ratchawan Ussama)

Candidate

ชื่อวิทยานิพนธ์	การประเมินความสามารถทักษะการสนทนา: การศึกษาเปรียบเทียบการประเมินแบบสนทนากับเพื่อน และแบบสัมภาษณ์ของนักเรียนไทยที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ
ผู้เขียน	นางสาวรัชต์วรรณ อุตมา
สาขาวิชา	การสอนภาษาอังกฤษเป็นภาษานานาชาติ
ปีการศึกษา	2555

บทคัดย่อ

งานวิจัยฉบับนี้ มีวัตถุประสงค์เพื่อศึกษาความแตกต่างของรูปแบบการประเมินทักษะการสนทนาสองรูปแบบ คือ แบบสนทนากับเพื่อน และแบบสัมภาษณ์ ของนักเรียนไทยที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ โดยมุ่งตอบคำถามวิจัยต่อไปนี้ 1) ผลการประเมินทักษะการพูดของแบบการประเมินทั้งสองรูปแบบแตกต่างกันหรือไม่อย่างไร เมื่อใช้ traditional rubric 2) ลักษณะของการโต้ตอบบทสนทนาของนักเรียนในรูปแบบการประเมินทั้งสองแบบนี้แตกต่างกันหรือไม่ และอย่างไร กลุ่มตัวอย่างที่ใช้เก็บข้อมูลเพื่อการวิจัยในครั้งนี้ประกอบด้วย นักเรียนชั้นมัธยมศึกษาปีที่ 3 จำนวน 10 คน ซึ่งเรียนในหลักสูตรที่มีการจัดการเรียนการสอนเป็นภาษาอังกฤษ นักเรียนกลุ่มนี้จัดว่าเป็นผู้ที่มีความสามารถทางภาษาอังกฤษในเกณฑ์สูง ผู้วิจัยให้นักเรียนทดสอบทักษะการพูดโดยใช้รูปแบบการประเมินทักษะการพูดที่แตกต่างกัน ตามหัวข้อที่ครูกำหนดให้ 3 หัวข้อ แบบแรกนักเรียนถูกสัมภาษณ์โดยครูผู้สอนชาวต่างชาติ ใช้เวลาในการทดสอบ 5 นาที และแบบที่สองนักเรียนคุยโต้ตอบกับเพื่อนในหัวข้อเดียวกันกับแบบสัมภาษณ์ โดยใช้เวลาในการทดสอบ 5 นาทีเช่นเดียวกัน โดยการทดสอบทั้งสองรูปแบบถูกบันทึกเทปเสียงบทสนทนาเพื่อการวิเคราะห์ข้อมูลตามหลัก Conversation Analysis (CA)

ผลการวิจัยพบว่าคะแนนของนักเรียนที่ได้จากการทดสอบทั้งสองรูปแบบไม่แตกต่างกันอย่างมีนัยสำคัญ อย่างไรก็ตามเมื่อเปรียบเทียบลักษณะการโต้ตอบของการสนทนาที่ปรากฏจากทั้งสองรูปแบบการประเมินมีความแตกต่างกัน ในประเด็น sequence opening, extending and closing, sequence abandonment, turn size, gesture-only turns, overlap และ repair initiation ซึ่งให้เห็นว่าการประเมินทักษะการพูดปฏิสัมพันธ์กับเพื่อน เปิดโอกาสให้ผู้เรียนได้ใช้ภาษาในลักษณะต่างๆ ที่พบในการสนทนาในชีวิตประจำวัน และสามารถวัดทักษะการสนทนาได้อย่างมีประสิทธิภาพมากกว่า

Thesis Title	Conversation Proficiency Assessment: A Comparative Study of Two-Party Peer Interaction and Interview Interaction Implemented with Thai EFL learners
Author	Ms. Ratchawan Ussama
Major Program	Teaching English as an International Language
Academic Year	2012

ABSTRACT

The study investigated the differences between two types of oral English proficiency tests, i.e., interview interaction and two-party peer interaction, in assessing learners' English speaking performance. It attempted to determine: a) whether the two types of assessment produce different results when used in assessing students' oral performance with a traditional rubric, and b) whether they differ in terms of the interactional features produced by the students. The participants of the study were ten students in the M.3 English Program at Pimanpittayasan School, Satun. Two test tasks were used: 5-minute interview interactions and 5-minute two-party peer interactions. The participants' conversations from both test tasks were videotaped, transcribed and analyzed according to Conversation Analysis (CA) principles.

The results revealed that the total scores obtained from both tasks using the traditional rubric were not significantly different at the .05 level. It was suggested that both interview and two-party peer interactions produce similar results when traditionally assessing the overall oral English performance of the learners. However, close analysis of the talks elicited from the two test tasks revealed differences in such interactional features as sequence opening, extending. This indicates that as far as conversational competencies are concerned, interview interaction may be of more limited value than peer interaction.

ACKNOWLEDGEMENTS

I acknowledge my deep gratitude towards my thesis advisor, Dr. Kemtong Sinwongsuwat, whose dedicated supervision, academic advice, constant encouragement and infinite patience have led to the successful completion of the thesis.

I would also like to note my gratitude towards Asst. Prof. Dr. Premin Karavi and Dr. Compol Swangboonsatic of the review committee for their valuable time and constructive comments. Additionally, I also want to express my gratitude towards Mr. Augustus R. Hall, who provided me with helpful comments and meticulously proofread my thesis.

My sincere gratitude also goes to all the M.A. instructors, as well as the M.A. staff of the Teaching English as an International Language program at the Department of Language and Linguistics, the Faculty of Liberal Arts, Prince of Songkla University, Hatyai Campus. In particular, I would like to thank Mr. Augustus R. Hall and Mrs. Sammireh Nagaratnam for sacrificing their time to rate the students' scores. In addition, I also want to extend my appreciation to fellow students in this program, for all their help, spiritual support, consistent assistance and constant encouragement.

It is my pleasure to convey my special thanks also to Pimanpittayasan School and to the 10 English Program students of M.3, who participated in this study. I would like to particularly thank Mr. Matthew Gallagher, an English Program teacher of Pimanpittayasan School, for his assistance in interviewing all the students for the data collection. Also, it is a pleasure to thank Mr. Anek Hirunyasatid, the head of the English Program at Pimanpittayasan School, for his help with data-collection facilities.

Words apparently fail me in expressing my profound gratitude to my beloved friend, Pichet Songka, for his beneficial advice, favorable suggestions, encouragement and support. I thank him very much for listening to my problems while writing the thesis. As well, my thanks go to all my beloved relatives, cousins,

and friends who always generously gave me their moral support and instilled learning inspiration in me. I was very fortunate to have them in my life.

Last but not least, my heartfelt thanks go to my dear father, mother and my beloved sister, who have always given their wholehearted care and constant encouragement, and kept on sponsoring me throughout my thesis.

Ratchawan Ussama

TABLE OF CONTENTS

ABSTRACT	v
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	ix
LIST OF TABLES	xi
LIST OF PAPERS	xii
LETTERS OF ACCEPTANCE	xiii
A SYNTHESIS REPORT	
1. INTRODUCTION	1
2. OBJECTIVES OF THE STUDY	6
3. RESEARCH QUESTIONS	7
4. RESEARCH METHODOLOGY	7
4.1 Participants of the study.....	7
4.2 Instruments.....	8
4.2.1 Test tasks.....	8
4.2.2 Testing Topics.....	8
4.2.3 Criteria for assessment.....	8
4.3 Data collection.....	9
4.3.1 Interview interaction	9
4.3.2 Two-party peer interaction	9
4.4 Data analysis.....	11
5. FINDINGS AND DISCUSSION	11
5.1 Results from the two test tasks used in assessing students’ oral performance with the traditional rubric.....	11
5.2 Different interactional features of students’ talk elicited by the two types of assessment.....	13
5.2.1 Sequence opening, extending, and closing.....	13
5.2.3 Sequence abandonment.....	14
5.2.4 Turn size.....	14
5.2.5 Gesture-only turns.....	14

TABLE OF CONTENTS (Continued)

5.2.6 Overlap.....	15
5.2.7 Repair initiation.....	15
5.3 Oral proficiency assessability of the two test tasks.....	15
6. CONCLUDING REMARKS AND RECOMMENDATIONS.....	16
REFERENCES.....	18
APPENDICES	
A: Scoring criteria and descriptors.....	20
B: Evaluation form.....	22
ACCEPTED MANUSCRIPT	
PAPER 1 Conversation Proficiency Assessment: A Comparative Study of Two-Party Peer Interaction and Interview Interaction Implemented with Thai EFL Learners	24
PAPER 2 A Comparative Study of Features of Talk Elicited by Two-Party Peer Interaction and Oral Interview Tests Implemented with Thai EFL Learners	37
VITAE.....	54

LIST OF TABLES

TABLE

A SYNTHESIS REPORT

1: Score results from two types of assessment	11
---	----

PAPER 1

1: Score results from two types of assessment.....	29
2: Interactional characteristics of the students' talks elicited from the two types of assessment tasks.....	30

LIST OF PAPERS

This thesis is based on the following papers:

1. Conversation Proficiency Assessment: A Comparative Study of Two-Party Peer Interaction and Interview Interaction Implemented with Thai EFL Learners
2. A Comparative Study of Features of Talk Elicited by Two-Party Peer Interaction and Oral Interview Tests Implemented with Thai EFL Learners

Dear Sinwongsawat,

The evaluation of your paper has started. Can you suggest any thai reviewers who are interested in TEFL?

Best Regards...

--

Dil ve Edebiyat Eđitimi Dergisi

(Uluslararası Online Hakemli Dergi)

Eriřim Adresi: <http://www.deedergisi.org/>

Journal of Language and Literature Education

(International online peer reviewed journal)

Website: <http://www.jllesite.org/>

ISSN: 2146-6971

--

Kemtong Sinwongsawat, Ph.D.

(English Language and Linguistics)

Dept. of Languages and Linguistics

Fac. of Liberal Arts, Prince of Songkla University-Hat Yai, Songkhla 90112

Thailand

O: 074-286-771

M: 84-859-3292



มหาวิทยาลัยหาดใหญ่
HATYAI UNIVERSITY

ม.หาดใหญ่ มหาวิทยาลัยเอกชนชั้นนำของเมืองไทย
รัฐคิด รัฐธรรม รัฐสำเร็จ

ที่ ม.ญ. 0499/2556

๓๐ พฤษภาคม 2556

เรื่อง การตอบรับการตีพิมพ์ผลงานวิชาการในรายงานสืบเนื่องจากการประชุม (Proceedings)

เรียน คุณรัชต์วรัณ ฤศมา

สิ่งที่ส่งมาด้วย รายชื่อคณะกรรมการพิจารณาผลงานวิชาการ จำนวน 1 ชุด

ตามที่มหาวิทยาลัยหาดใหญ่ ได้ดำเนินการจัดการประชุมหาดใหญ่วิชาการ ครั้งที่ 4 เรื่อง “การวิจัยเพื่อพัฒนาสังคมไทย” โดยมีวัตถุประสงค์เพื่อให้เป็นเวทีทางวิชาการให้กับอาจารย์ นักวิจัย นักวิชาการ และนักศึกษาระดับบัณฑิตศึกษาได้มีโอกาสเผยแพร่ผลงานวิชาการและแลกเปลี่ยนองค์ความรู้และประสบการณ์ด้านการวิจัย และเป็นการสร้างบรรยากาศการวิจัยให้เกิดขึ้นในมหาวิทยาลัย ในวันศุกร์ที่ 10 พฤษภาคม 2556 เวลา 08.30 – 16.30 น. ณ อาคารศูนย์กีฬาและกิจการนักศึกษา มหาวิทยาลัยหาดใหญ่ อำเภอหาดใหญ่ จังหวัดสงขลา และท่านได้เข้าร่วมนำเสนอผลงานวิชาการแบบ oral presentation นั้น

ในการนี้ มหาวิทยาลัยหาดใหญ่ ขอเรียนให้ทราบว่าขณะนี้ผลงานทางวิชาการของท่านได้ผ่านการพิจารณาจากผู้ทรงคุณวุฒิในสาขาวิชา และยอมรับให้ตีพิมพ์ผลงานวิชาการของท่าน ในรายงานสืบเนื่องจากการประชุมหาดใหญ่วิชาการ ครั้งที่ 4 (Proceedings) และคาดว่าจะแล้วเสร็จสามารถเผยแพร่ไปยังหน่วยงานต่างๆ ได้ภายในเดือนสิงหาคม 2556

จึงเรียนมาเพื่อทราบ

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.วิวัฒน์ จันทร์กึ่งทอง)

ผู้อำนวยการสำนักวิจัยและพัฒนา

ปฏิบัติการแทนรองอธิการบดีฝ่ายวิจัยมหาวิทยาลัยหาดใหญ่

สำนักวิจัยและพัฒนา

โทรศัพท์ 0-7420-0300 ต่อ 108

โทรสาร 0-7420-0384

Hatyai University, the First Private University in Southern Thailand

125 / 502 Polpichai Rd. Hatyai Songkhla 90110 Thailand <http://www.hu.ac.th>
Tel 66 74 200300, 425000 Fax 66 74 425467

1. INTRODUCTION

Speaking skills are an essential part of any English language teaching curriculum, thus being an important object of assessment. The common purposes of a speaking assessment are to measure language proficiency; to assess achievement of the objective of a course; to diagnose learners' strengths and weaknesses; to identify what they do and do not know; and to assist in the placement of learners in a teaching program (Hughes, 2003). When assessing conversing ability, it is important to design a test that allows candidates to demonstrate their ability to use language in ways which are characteristic of their interactional competence. To test whether learners can converse, it is necessary to get them to take part directly in speaking activities. Thus, the oral tasks should elicit behavior which is valid, truly represents the learners' conversation competencies, and can be reliably scored.

In an EFL classroom, there are several types of speaking test tasks that are commonly used to assess learners' speaking ability. The formats available range from the more direct types such as interaction between students and face-to-face interview to the more indirect types such as response elicitation. Interview, however, is one of the most popular means of testing speaking skills (Weir, 1993). It is a direct, face-to-face exchange between learner and interviewer in which a learner's performance is evaluated. Generally, the interview tasks implemented in the class room are of two types: the unstructured interview and the controlled interview. The unstructured interview is like extended conversation, and the direction is allowed to unfold as the interview takes place. In the controlled interview, on the other hand, the interviewer normally has decided on the questions to ask to find out about the learner's language ability. The interviewer normally manages the interaction and retains the initiative in selecting and developing the topics. There is sometimes an assessor present who does not take part in the spoken interaction but listens, watches, and evaluates the abilities of the learner (Weir, 1993).

The advantages of the interview assessment were noted by Peace Corps (2005)--a unit of the office of overseas programming and training support in

United State. Namely, language proficiency interview provides a language to describe linguistic competencies and therefore determines what speakers can do with what they have learnt, and how effectively they can communicate with native speakers. This occurs on the condition that the interview is not a set of prearranged questions but a conversation in which the interviewer adapts and changes topics and functions in order to obtain a natural target language.

While widely practiced in assessing learners' oral ability, interviews apparently exhibit unequal distribution of conversation rights and duties between the participants. Thus, there have been mounting criticisms against the use of interviews in speaking assessment in recent years. Many researchers have in fact come to agree that the oral exchange that occurs between an interviewer and a test taker does not reflect or even closely replicate natural or real-life conversation (e.g. Bachman, 1990; Lazaraton, 1992; Van Lier, 1989; Young, 1995). Interview tests often resulted in a *test discourse* or an *institutional talk* that neither represented normal conversation nor provided candidates with the opportunity to show their ability to participate in interaction other than as an *interviewee* responding to questions.

Moreover, when testing oral performance and assessing interactional competence, the main concern is to uncover how speakers structure and sequence their speech and how they apply turn-taking rules, but interview tests make it difficult to achieve this. The oral interview led to an increase in empirical studies of the nature of its discourse during the 1990s. It was found that interview discourse was characterized by a power differential, turn-taking, topic organization, sequence and overall structure being predetermined or controlled by the interviewer (Lazaraton, 1992; Young & Milanovic, 1992).

Consequently, since the late 1980s, as one of the most common practices in classroom activities, pair or group tasks have increasingly been used to assess speaking ability (Egyud & Glover, 2001). In these tasks, test candidates work in pairs or groups while the examiner takes no part; the candidates should be more at ease and they have more opportunity and inclination to speak. From a pragmatic perspective, pair or group assessment is typically also more time- and cost- efficient

as learners are tested together, and raters assess two or more candidates simultaneously. This type of speaking test can provide opportunities for candidates to display not only their communicative competence but also a range of conversational management skills. Additionally, using pair tasks avoids several of the criticisms associated with an interview task, such as the power differential between speakers, and a question-and-answer style of discourse that may not very well reflect actual conversation (Young & Milanovic, 1992).

As far as the interactional features of oral performance are concerned, Luk (2010) recently attempted to investigate students' discourse performance in L2 oral proficiency assessment conducted in the form of peer group interaction in Hong Kong. Forty-three female Hong Kong secondary students were involved. Finding from a qualitative discourse analysis revealed that the turn organization appeared when students engaged themselves in procedural matters of talk, for example, opening and closing the talk, distributing the turn to the next speaker, or directing the talk to different topics. Additionally, not only such features occurred in the peer group task but the speech acts such as informative, reasoning, and hypothesis also emerged as students responded to peers' comments, questions, and requests for clarifications. These findings seem to indicate that pair- and group-format speaking tests provide opportunities for students to help each other to maintain the talk-in-interaction which taps into their ability to conduct ordinary conversation.

The advantages of the paired format of the speaking test were also described in Saville & Hargreaves (1999). The authors revised the First Certificate of English (FCE) examination produced by the University of Cambridge Local Examinations Syndicate (UCLES), and provided reasons for making decision to use the paired format as the standard model for the main suite speaking test. They emphasized the following advantages of the paired format; namely, the candidates are more relaxed, and more likely to vary patterns of interaction during the test. Another alleged advantage of this format is that it gives candidates a better chance to perform what they can do in the target language (Foot, 1999).

To sum up, based on the advantages of the paired test format discussed, it is evident that when students engage in the oral paired assessment, several salient features of talk-in-conversation inevitably emerge. Moreover, such a format not only provides the potential for a wider range of interactional moves but apparently reflects features of ordinary, naturally-occurring conversation.

Apart from the test formats eliciting different interactional features, another equally important factor contributing to the validity of the direct oral assessment is the criteria or scoring rubric used. Often accompanied by a scoring guide, a scoring rubric is a piece of paper that contains scoring scales provided with descriptors, or scoring criteria, the purpose of which is to describe briefly what the typical learner at each level can do (Lazaraton, 2002). According to Underhill (1987), the rubric widely used for rating students' oral performance often concentrates on the language produced, reflecting the view that the accurate command of language is an end in itself. Familiar components of language proficiency referred to in such a rubric, henceforth *the traditional rubric*, usually are grammar, vocabulary, pronunciation, style, fluency and content. In each of these categories, the rater might score on the basis of impression, or there might be a separate subsystem of scoring.

However, in recent decades, given the shift in emphasis from language as the end product to language as a tool for communication or as an action in interaction, the scoring rubric developed has started to look different. Apparently more suitable for assessing interactive communication, the more recent rubric covers various aspects of speakers' performances, for example:

- Size (how long are the utterances produced?)
- Complexity (how far does the speaker attempt?)
- Speed (how fast does he/she speak?)
- Flexibility (can the speaker adapt quickly to changes in the topic or task?)
- Accuracy (is it correct English?)
- Appropriacy (is the style or register appropriate?)

- Independence (does the speaker rely on a question or stimulus, or can he/she initiate speech on his own)
- Repetition (how often does the question or stimulus have to be repeated)
- Hesitation (how much does the speaker hesitate before and while speaking)

(Underhill, 1987, p. 96).

Additionally, a more skill- or strategy-oriented rubric has also been proposed based on the suggestion that students' oral performance be assessed according to how well they can use the skill and strategies that the assessing activity requires (Luoma, 2005). For instance, in information-related talk such as explanations or narratives, success partly depends on the content and sequencing of the examinees' talk. Therefore, the rubric should contain items such as information sequencing or event and action ordering.

While there have been questions regarding the appropriateness of the interview test format and the traditional rubric used in assessing students' conversation proficiency, it is still too early to lay claim that either the interview interaction or the two-party peer interaction better assesses Thai EFL learners' conversing ability. In fact, there have been only few studies directly comparing the results from both oral proficiency test tasks used with the traditional rubric in assessing the learners. Neither have there been any studies which compared interactional features of the learners' talks elicited from the two test tasks. Therefore, this study was designed to fill this gap by investigating the differences between the two tests in terms of the overall oral performance traditionally assessed, its discrete aspects, and the interactional features shaped by the tests.

1.1 Definitions of key terms

1.1.1 Oral proficiency assessment refers to the direct test to determine whether the candidates can speak and to assess their conversing performance.

1.1.2 Interview interaction refers to the unstructured interview in which the student responds to an interviewer who is a native speaker teacher.

1.1.3 Two-party peer interaction refers to the test task in which the students interact with their peer.

1.1.4. Traditional rubric refers to the oral-performance scoring rubric with criteria oriented towards features of linguistic competencies such as fluency, vocabulary, appropriacy, comprehensibility, and grammar.

1.1.5 Conversation competency is the ability to perform openings and closings, to establish and change topics, to hold and yield the floor, to interrupt and to collaborate, as well as to recognize and produce adjacency pairs.

1.1.6 Thai EFL learners refer to the Thai students who had attended the English Program at Phimanpittayasan School, Satun Province.

2. OBJECTIVES OF THE STUDY

The main purposes of this study are:

2.1 to investigate whether the two types of assessment; i.e., interview interaction and two-party peer interaction, produced different results when used in assessing the students' oral performance with the traditional rubric.

2.2 to determine whether the two types of assessment differed in terms of the interactional features produced by the students.

It was hoped that the study could shed light on how the differences in elicited interactional features possibly contribute to the different score results and the conversational assessability of each test task. Additionally, it would allow teachers to make an informed decision on which oral test tasks to use when assessing the students' English conversational competence. It was also expected to help the teachers

to establish criteria which include the important features of talk suitable for each test task.

3. RESEARCH QUESTIONS

The following two research questions were addressed accordingly:

3.1 Do the interview interaction and the two-party peer interaction produce different results when used in assessing students' oral performance with the traditional rubric? If so, how?

3.2 Do the two types of assessment differ in terms of the interactional features produced by the students? If so, how?

4. RESEARCH METHODOLOGY

4.1 Participants of the study

The participants of this study were ten high-proficiency students sampled from the total population of 32 students who were in M.3, or Grade 9, in the English Program at Phimanpittayasan School. They had taken a Listening and Speaking course in the first semester of the academic year 2011 and took the oral tests as a part of the end-of-the year examination. The class was divided into three proficiency levels; i.e., high, middle, and low, according to the grade of a compulsory English course of the academic year 2010.

Only the high level group was chosen for data collection and analysis for the reason that the high proficiency students seem to produce more talk when working with partners. This was based on the belief that the interlocutor proficiency may influence both the amount of talk and scores received in a speaking task (Iwashita, 1996). Thus, if interlocutor proficiency influences the amount of language an examinee produces, it seems reasonable to suspect that the nature of the interaction, and ultimately the scores might be influenced as well.

4.2 Instruments

The sets of materials involved in the data collection of this study include the test tasks, topic used in testing, and criteria for assessment.

4.2.1 Test tasks

The two test tasks used were:

- (1) a 5-minute interview interaction in which a native speaker interviewer asked questions related to the topic that the students had prepared, and
- (2) a 5-minute two-party peer interaction in which two test candidates talked about the same topic as they did in the interview interaction.

4.2.2 Testing Topics

The three topics used in the two test tasks were assigned by the teacher before testing. The topics chosen were related to the content of the course as follows:

- (1) What is tourism?
- (2) What do you like and dislike about tourism in Thailand?
- (3) As you are local people in Satun, which places would you like to take the tourists to visit? Why?

4.2.3 Criteria for assessment

In this study the students' oral performance was rated by two assessors who are native speakers of English and did not take part in the test. A traditional interview-scoring rubric based on Underhill (1987) and Lazaraton (2002) was deployed for rating the students' speaking proficiency. The raters were given not only the rubric along with a descriptive set of rating topics including fluency, vocabulary, appropriacy, comprehensibility, and grammar (See Appendix A), but also the evaluation form to evaluate individual students' performance (See Appendix B). Each topic was rated on a 5-point scale: 1= very poor/ unacceptable, 2= poor, 3= average, 4= good, 5= excellent.

4.3 Data collection

The data was collected during the first semester of the 2011 academic year from July to September. In this study, the oral assessment was used as the achievement test of the course. That is, the two test tasks were conducted to assess the students' oral performance after the course completion. They were administered three times each on different topics. The students were assigned to research the topic before the test task was conducted. Generally, the three topics were introduced in both tasks in the following sequence: *What is tourism?; What do you like and dislike about tourism in Thailand?; and As you are local people in Satun, which places would you like to take the tourists to visit? Why?*. The candidates were first engaged in an oral interview and subsequently during the following week in the two-party peer interaction on the same topic. The students took the test in an English laboratory, and the performances of both tasks were videotaped.

4.3.1 Interview interaction

The students were allowed to do research on the topic required by the teacher before doing the test. In the interview interaction, the interviewer (their teacher) started the question by discussing 2-3 everyday topics and subsequently asking about the information on the topic prepared. Each interview took approximately five minutes.

4.3.2 Two-party peer interaction

In the two-party peer interaction assessment, the peers were matched by the teacher and each pair was randomly asked to undertake the test task. This task assessed the students on the same topic as the interview. The students were asked to share the prepared information with each other. The duration of the test was approximately five minutes. When finished with the pair task, the students were asked to prepare for the interview task of the following week.

Both of the oral test tasks were video-recorded and scored by two native raters who did not take part in the oral assessment according to the criteria previously mentioned. The conversations were transcribed and analyzed using the following transcription convention adopted by Seedhouse (2004) and Schegloff (2007).

[Point of overlap onset
]	Point of overlap termination
=	(a) Turn continues below, at the next identical symbol (b) If inserted at the end of one speaker's adjacent turn, indicates that there is no gap at all between the two turns (c) Indicates that there is no interval between adjacent utterances
(0.5)	Numbers in parentheses indicate silence, represented in tenths of a second; what is given here indicates 5 seconds of silence
(.)	Very short untimed pause; ordinarily less than 2 seconds
<u>word</u>	Speaker emphasis
-	A hyphen after a word or part of a word indicates a cut-off or self-interruption
?	Rising intonation, not necessarily a question
.	Low-rising intonation, or final, not necessarily the end of a sentence
()	A stretch of unclear or unintelligible speech
wo:rd	Colons show that speaker has stretched the preceding sound
◦word◦	Material between "degree signs" is quieter than the surrounding talk
((word))	Transcriber's comments
[gibee]	In the case of inaccurate pronunciation of an English word, an approximation of the sound is given in square brackets
<i>ja</i> ((tr.: yes))	Non-English words are italicized and followed by an English translation in double parentheses

→ Mark features of special interest

4.4 Data Analysis

To answer the two research questions above, the scores obtained from the two raters in each test task were averaged by using mean standard variation and the *t*-test was carried out to determine the significant degree of difference in the students' oral performance. The transcribed data were closely examined with reference to different interactional features oriented to by the talk participants in the two oral test tasks.

5. FINDINGS AND DISCUSSION

5.1 Results from the two test tasks used in assessing students' oral performance with the traditional rubric

Results from the interview and the two-party peer interaction test tasks used with the traditional rubric are displayed and discussed below.

Table 1: *Score results from the two types of assessment*

Criterion	Interview interaction		Two-Party peer interaction		<i>t</i> -value	Sig.
	\bar{x}	S.D.	\bar{x}	S.D.		
Fluency	9.25	1.36	9.15	1.38	0.28	0.4
Vocabulary	8.9	1.2	9.2	0.71	-0.94	0.19
Appropriacy	9.75	1.14	10	0.88	-0.6	0.28
Comprehensibility	9.9	1.41	10.05	1.32	-0.34	0.37
Grammar	8.45	0.98	8.95	0.6	-1.94	0.04*
Total	46.25	5.28	47.35	4.24	-0.83	0.22

* Significant at 0.05 level

As shown in Table 1, based on the *t*-test, the total scores obtained from both test tasks were not significantly different at the 0.05 level ($t = -0.83$, $p = 0.22$). It was suggested that both interview interaction and two-party peer interaction

produced similar results overall when used in assessing the learners' conversational performance with the traditional linguistic-competency-oriented rubric. Interestingly, although not statistically different in the scores, the students appeared to perform better in the peer interaction. The mean scores on discrete items, e.g., vocabulary, appropriacy, and comprehensibility of peer interaction were higher than those obtained from the interview interaction. The higher mean scores on these aspects in peer interaction may be due to two facts peculiar to the format. First, its looser structure provided more opportunities for the students to use the full range of their English. In other words, the peer interaction increased opportunities for the test takers to display their conversational management skills, and the students were therefore positively received by the raters. Moreover, two-party peer interaction by definition is between subjects of similar if not the same background, social position, etc. In this case, the students were from the same school and class. Therefore, prior knowledge of an interaction with said peers made for easier communication and greater intersubjectivity. They not only were more relaxed but also could perceive and even predict their interlocutor understanding or lack thereof from shared experiences, visual cues, etc.

Additionally, it was particularly interesting to note that in terms of grammar, the students scored significantly better in their interaction with peers than with the native speaker. This indicated that students were more conscious of their grammar talking with a peer. This may seem counter intuitive, but the students seemingly knew that the native speaker interviewer (their teacher) would be more tolerant and understand them regardless of their grammatical errors. On the other hand, with a peer interlocutor, knowing that they were being observed by a third party, the researcher, and probably fearing communication misunderstanding and losing face, they made an effort to produce grammatically accurate turns and often initiated self-repair on grammar.

As far as fluency is concerned, the students however scored better when interacting with the native speaker interviewer than with their peers, suggesting that the interviewer teacher, who played a dominant role in the interaction, appeared

to be more effective in noticing and dealing with problematic, dispreferred turns. Through sequence opening and closing turns, topic/sequence extension, collaborative turn completion, repair initiation by asking for clarification, seeking confirmation, and supplying correct forms of words, gaps in and between turns seemed to be reduced, making the conversation sound more fluent. Therefore, in line with Davis (2009), it can be implied that that lower-proficiency speakers may benefit from working with an interlocutor who is more conversationally competent in the target language.

5.2 Different interactional features of students' talk elicited by the two types of assessment

Different interactional features were revealed in the talks obtained from both test tasks.

5.2.1 Sequence opening, extension, and closing (See Paper 2, p. 43 for illustrations)

In the interview interaction, since it was the interviewer who had rights to open and close the dialogue and to ask questions and introduce a new topic, the students did not get to perform any of these actions. They were mostly prompted to answer the questions asked by the interviewer (their teacher), and were not given much opportunity to negotiate for meaning as the teacher played a dominant role in turn initiation and allocation. The students in two-party peer interaction, on the other hand, made a collaborative effort to construct turns and maintain ongoing sequences. They opened, extended, and closed the sequences all by themselves. In line with Kormos (1999), through such interactional features as sequence openings and closings, introducing new topics and topic shift, they were apparently able to exercise their most basic rights in conversation.

Moreover, the students engaged in the interview interaction task had no opportunity to establish or initiate the topic of the talk. It was the interviewer who always proffered the topic via open and closed questions to which the students provided responses by recycling and topicalizing the information provided by the

interviewee. In contrast, in the two-party interaction, the students could balance both roles--topic initiators and responders--and participate in a more reciprocal exchange.

5.2.2 Sequence abandonment (See Paper 2, p. 46 for illustrations)

The students engaged in the interview interaction additionally had little chance to reformulate trouble turns since most of the turns were repaired by the interviewer. Taking the dominant role, the teacher completely controlled the direction of the sequence. It seemed that such conversations were performed fluently though the answer was omitted, thus abandoning the sequence. When interacting with peers, they showed an attempt to repair the trouble turns or shift the topic if the attempt failed.

5.2.3 Turn size (See Paper 2, p. 46 for illustrations)

Concerning turn size, in the interview assessment, since the abilities to initiate or extend the topic were controlled by the interviewer, the students' turns seemed to be smaller and less complex than the interviewer's. In the interview interaction, the teacher obviously emerged as the more dominating speaker as she spoke more, took longer turns, made all the initiation moves, and was the only speaker to expand on her ideas. In contrast, the quantity of the talk was balanced between the two participants in peer tasks. The rights and duties of the interactants were apparently equally distributed in the speakers' turns.

5.2.4 Gesture-only turns (See Paper 2, P. 47 for illustrations)

The students in interview interaction also employed more gesturing turns. When confronted with difficult questions asked by the teacher, they tried to respond to the question first with gestures and later with words even though they mostly failed to produce a complete turn, ending up with a smile and a pause. Similarly, struggling with certain words, they also resorted to gestures to mime the target word in the peer interaction.

5.2.5 Overlap (See Paper 2, pp. 48-49 for illustrations)

The talks elicited from the two test tasks also differed as far as overlaps are concerned. In the case of two-party peer interaction, the frequent

overlaps did not result in a topic shift but an extension of the prior topic or a support for the speaker who initiated the topic. On the other hand, in the interview interaction, the student proffered only a minimal response such as “yes” or nods with no overlap.

5.2.6 Repair initiation (See Paper 2, pp. 49-51 for illustrations)

Another different salient feature of the interaction elicited from the two test tasks lies in repair organization. The data transcribed showed that the important type of repair organization occurred in both test tasks was repair for meaning and fluency. Linguistic errors were often ignored in both test tasks if they did not obscure the meaning. The repair strategies used varied depending on the tasks. Collaborative completion was not found at all in the interview task while a wider range of repair strategies were resorted to in the two-party peer interaction task such as collaborative completion, asking for repetition, asking for clarification and seeking confirmation.

5.3 Oral proficiency assessability of the two test tasks

Based on the findings, it can be suggested that either interview or peer-peer interaction test task can be used to assess students' linguistic competency in oral communication since they do not produce significantly different results when used with the traditional rubric. However, as far as conversational competencies are concerned, the interview task may be of limited value since it does not allow the students to produce the interactional features sufficiently characterizing their conversational competence, which is the crucial component of EFL oral proficiency.

Being an unequal social encounter, the interview task does not create a situation in which natural conversations can occur. The two-party peer interaction, on the other hand, provides the test candidates more opportunities to perform relevant interactional features at talk, particularly those related to organizing adjacency pair sequences, openings, closings, initiating, extending, and shifting topics, as well as dealing with problems in talk.

6. CONCLUDING REMARKS AND RECOMMENDATIONS

This study examined whether two types of assessment, i.e., interview and two-party peer interactions, produced different results when used in assessing students' oral performance with a traditional rubric. On the basis of the analysis of 10 Thai EFL students' performance through the two test tasks, i.e., interview interaction and two-party peer interaction, it can be concluded that using the traditional rubric, the total scores obtained from the two test tasks were not significantly different even though the students in the peer interaction task performed better in most of the discrete items especially in terms of grammar. These may be due to the fact that the students were provided more opportunities to formulate their own turns, direct and maintain the conversation themselves. Additionally, they apparently felt more relaxed when conversing with their peer.

The results of the analysis of how the two types of assessment differ in terms of the interactional features elicited suggested that the interview interaction may not be appropriate for assessing students' conversational competencies. Due to the unequal distribution of power, candidates in the interview interaction rarely have the chance to initiate a new topic and have no rights to initiate the opening or closing of the conversation, which are basic to conversational ability. In contrast, peer interaction is recommended as a better means of assessing candidates' conversational competence.

The findings of this study could pose recommendations for EFL educators and further research as follows.

In order to determine the reliability of the different test tasks used for measuring students' oral proficiency, there is a need for more controlled studies. In such studies, the 1-5 rating scale should be made clear to the raters by specifying correspondent speech features. Additionally, even though unstructured peer interview tasks may be able to elicit more features of natural conversation, future studies should select familiar topics which interest the participants and employ such tasks as non-scripted role-play to elicit their talk. In role-play activities, participants are given more

freedom to initiate the conversation with their own turns without having to be concerned with the list of prescribed questions and are constrained only by specific roles and speech situations, more approximating natural conversation.

Furthermore, rather than focusing primarily on linguistic competencies, future studies which desire to assess learners' conversational competencies with more accuracy should also consider using a rubric which includes such interactional features as reported. It would be worthwhile to examine how learners of different levels of English proficiency perform in the task using such a rubric.

This study has not only shed light on the design of more appropriate speaking tasks and rubrics to assess learners' conversational competencies but also helped raise teachers' awareness of the interactional features common to natural conversations, which really need to be taken into account if we want to make the prevalent oral proficiency test more valid.

References

- Bachman, L. (1990). *Fundamental consideration in language testing*. Oxford: Oxford University Press.
- Davis, L. (2009). The influence of interlocutor proficiency in a pair oral assessment. *Language Testing*, 26, 367-396.
- Egyud, G., & Glover, P. (2001). Oral testing in pairs – A secondary school perspective. *ELT Journal*, 55(1), 70-76.
- Foot, M.C. (1999). Relaxing in pairs. *ELT Journal*, 53(1), 36-41.
- Hughes, A. (2003). *Testing for language teacher*. Cambridge: Cambridge University Press.
- Iwashita, N. (1996). The validity of the paired interview format in oral performance assessment. *Melbourne Paper in Language Testing*, 5, 1-65.
- Kormos, J. (1999). Simulating conversations in oral proficiency assessment: A conversation analysis of role plays and non-scripted interview in language exams. *Language Testing*, 16(2), 163-188.
- Lazaraton, A. (1992). The structural organization of a language interview: A conversation analytic perspective. *System*, 20, 373-86.
- Lazaraton, A. (2002). *A qualitative approach to the validation of oral language tests*. Cambridge: Cambridge University Press.
- Luk, J. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly*, 7, 25-53.
- Luoma, S. (2005). *Assessing speaking*. Cambridge University Press.
- The Peace Corps Information Collection and Exchange. (2005). *Language proficiency interview (Manual for testers)*. Washington, DC: ICE Press.
- Saville, N., & Hargreaves, P. (1999). Assessing speaking in the revised FCE. *ELT Journal*, 53(1), 42-51.
- Schegloff, E.A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press.

- Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. United Kingdom: Blackwell Publishing.
- Van Lier, L. (1989). Reeling, writhing, drawling, stretching and fainting in coils; oral proficiency interviews as conversation. *TESOL Quarterly*, 23, 489-508.
- Underhill, N. (1987). *Testing spoken language*. Cambridge: Cambridge University Press.
- Weir, C.J., (1993). *Understanding and developing language tests*. United Kingdom: Prentice Hall.
- Young, R. (1995). Conversational styles in language proficiency interviews. *Language Learning*, 45, 3-42.
- Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. *Study in Second Language Acquisition*, 14, 403-424.

Appendix A
Scoring criteria and descriptors

Scoring criteria and descriptors adapted from Lazaraton (2002) and Underhill (1987)

Fluency	Vocabulary	Appropriacy	Comprehensibility	Grammar
Students will be able to speak fluently with only occasional hesitation. Speech may be slightly slower than that of a native speaker.	Students will be able to use vocabulary broad enough to allow the candidate to express most ideas well.	Students will be able to demonstrate good awareness of social conventions and have some abilities to respond to unpredictable turns in conversation, though may sound unnatural.	Students will be able to understand most speech or may require repetition of details.	Students will be able to generally demonstrate control of varieties of structures with only minor errors.

Appendix B
Evaluation form

Evaluation form

Name.....

Topic.....

Score Topic	Excellent 5	Good 4	Satisfactory 3	Poor 2	Very poor/ unacceptable 1
Fluency					
Vocabulary					
Appropriacy					
Comprehensibility					
Grammar					

Comments:.....

PAPER 1

Conversation Proficiency Assessment: A Comparative Study of
Two-Party Peer Interaction and Interview Interaction
Implemented with Thai EFL Learners

Conversation Proficiency Assessment: A Comparative Study of Two-Party Peer Interaction and Interview Interaction Implemented with Thai EFL Learners

Ratchawan Ussama¹
Kemtong Sinwongsuwat, Ph.D.²

1. Graduate Student, Master of Arts in Teaching English as an International Language, Faculty of Liberal Arts, Prince of Songkla University, Thailand, angelicasul@windowslive.com
2. Lecturer, Department of Language and Linguistics, Faculty of Liberal Arts, Prince of Songkla University, Thailand, ksinwong@gmail.com

Abstract

The study investigated the differences between two types of oral English proficiency tests, i.e., interview interaction and two-party peer interaction, in assessing learners' English speaking performance. It attempted to determine: a) whether the two test types produce different results when used in assessing students' oral performance with a traditional rubric; b) whether they differ in terms of the interactional features produced by the students; and c) how the differences in interactional features possibly contribute to the different score results and the conversation-assessing ability of each test task. The participants of the study were ten students in the Grade 9 English Program at Pimanpittayasan School, Satun, Thailand. Two test tasks were used: 5-minute interview interactions and 5-minute two-party peer interactions. The conversations obtained from both test tasks were videotaped, transcribed and analyzed according to Conversation Analysis (CA) principles. The results revealed that the total scores obtained from both tests using the traditional rubric were not significantly different at the .05 level. It was suggested that both interview and two-party peer interactions produced similar results when traditionally assessing the overall oral English performance of the learners. However, close analysis of talks elicited from the two test tasks revealed different interactional features some of which were related to learners' conversational problems, indicating that as far as conversational competencies are concerned, interview interaction may be of more limited value than peer interaction.

Keywords: Oral proficiency assessment; interview interaction; two-party peer interaction; Conversation Analysis (CA); linguistic competencies; conversational competencies; Thai EFL learners

Introduction

Speaking skills are an essential part of any English language teaching curriculum, thus being an important object of assessment. The common purposes of a speaking assessment are to measure language proficiency; to assess achievement of the objective of a course; to diagnose learners' strengths and weaknesses; to identify what they do and do not know; and to assist in the placement of learners in a teaching program (Hughes, 2003). When assessing conversing ability, it is important to design a test that allows candidates to demonstrate their ability to use language in the ways which are characteristic of their interactional competence. To test whether learners can converse, it is necessary to get them to take part directly in speaking activities.

In an EFL classroom, there are several types of speaking test tasks that are commonly used to assess learners' speaking ability. An interview is one of the most popular means of testing speaking skills (Underhill, 1998; Weir, 1993). It is a direct, face-to-face exchange between a learner and interviewer in which a learner's performance is evaluated. There is sometimes an assessor present who does not take part in the spoken interaction but listens, watches, and evaluates the abilities of the learner (Weir, 1993).

While widely practiced, there has been mounting criticism against the use of interviews in speaking assessment. Many researchers have come to agree that the oral exchange that occurs between an interviewer and a test taker does not reflect or even closely replicate natural or real-life conversation (e.g. Bachman, 1990; Lazaraton, 1992; Van Lier, 1989; Young, 1995). Interview tests often result in a test discourse or an institutional talk that neither represents normal conversation nor provides candidates with the opportunity to show their ability to participate in interaction other than as an interviewee responding to questions.

Consequently, since the late 1980s, as one of the most common practices in classroom activities, pair or group tasks have increasingly been used to assess speaking ability (Egyud & Glover, 2001). From a pragmatic perspective, pair or group assessment is typically also more time- and cost-efficient as learners are tested together, and raters assess two or more candidates simultaneously.

Examining raters' orientation in oral assessment, Ducasse and Brown (2009) reported that interactional and conversational management was the function of language use oriented to by the raters in paired tasks. The manner in which turns are organized verbally and nonverbally was recognized as the features contributing to successful interaction, thus subject to being assessed. The paired test format made possible the assessment of several salient features of talk and provided the potential for a wider range of interactional moves apparently reflecting genuine features of ordinary conversation.

Despite the argument that the interview was not as appropriate as paired peer interaction for measuring conversation skills in oral tests, it is too early to claim that either the interview interaction or the two-party peer interaction can better assess Thai EFL learners' oral performance. In fact, there have been only few studies directly comparing the results of both oral proficiency test tasks used with the traditional rubric oriented towards such features of linguistic competencies as fluency, vocabulary, appropriacy, comprehensibility, and grammar. Additionally, there have been no studies that compared interactional features of Thai learners' talks elicited from the two test tasks. Therefore, this study is designed to fill this gap by investigating the differences between the two tests in terms of the overall oral performance traditionally assessed, its discrete aspects and the interactional features shaped by the tests.

Three research questions explored in the present study are listed below:

- (1) Do the two types of assessment produce different results when used in assessing Thai students' oral performance with the traditional rubric? If so, how?
- (2) Do the two types of assessment differ in terms of the interactional features produced by the students? If so, how?
- (3) How do the differences in interactional features contribute to the conversation proficiency-assessing ability of each test task?

Significance of the study

This study sheds light on the differences of the two types of oral English proficiency tests in terms of the overall oral performance traditionally assessed, its discrete aspects and the interactional features of the talks induced. It should aid the teacher in making a decision on which oral test tasks to use when assessing the students' English conversational proficiency.

Research methods

In this study, the oral assessment was used as an achievement test to assess the students' oral performance during the course. Each test was conducted three times over the course of three months. The learners were first engaged in an oral interview interaction and at a later date in a two-party peer interaction on the same topic.

Prior to the interview interaction, the students were allowed to do research on the topic assigned by the teacher. The interviewer started the questions by discussing every day topics and subsequently inquiring about the information the students had prepared. Each interview took approximately 5 minutes. In the two-party interaction assessment, on the other hand, the student peers were matched by the teacher, and just as in the interview, the students were asked to discuss the topic prepared for about 5 minutes. The interaction was assessed on the same criteria as the interview interaction.

Both of the oral test tasks were video-recorded and scored by the assessors according to the criteria including features such as fluency, vocabulary, appropriacy, comprehensibility, and grammar. The transcription of the recorded data was carried out by the researcher for subsequent close analysis.

Population and sample

The participants in this study were ten students sampled from the total population of 32 students who were in the Grade 9 English Program at Pimanpittayasan School, Satun Province, Thailand. They attended a Listening and Speaking course in the first semester of the academic year 2011 and took the oral tests as a part of the end-of-the year examination. The class was divided into three levels based on their language proficiency: high, middle, and low according to their grade in a compulsory English course of the academic year 2010. Only the high level group was chosen for the investigation using the 33% technique due to the belief that high proficiency students would produce more talk to be assessed when working with partners (Iwashita, 1996).

Instruments

Test tasks

The two test tasks used were:

- (a) a 5-minute interview interaction in which a native speaker interviewer asked questions related to the topic that the students had prepared, and
- (b) a 5-minute two-party peer interaction in which two student candidates talked about the same topic as the interview interaction.

The three topics used in the two test tasks were assigned by the teacher before testing. The topics chosen were related to the content of the course as follows: “What is tourism?”; “What do you like and dislike about tourism in Thailand?”; and “As you are local people in Satun, which places would you like to take the tourists to? Why?”

Criteria for assessment

In this study, the students’ oral performance was rated by two assessors who are native speakers of English and did not take part in the test. A traditional interview-scoring rubric based on O’Loughlin (2001) was deployed for rating the students’ speaking proficiency. The raters were given the rubric along with a descriptive set of rating topics including fluency, vocabulary, appropriacy, comprehensibility, and grammar (See Appendix 1). Each topic was rated on a 5-point scale: 1= very poor/ unacceptable, 2= poor, 3= average, 4= good, 5= excellent. Then, the inter-rater reliability scores were computed to ensure the similarity in the scores of the two independent assessors.

Data analysis

To answer the three research questions above, the scores obtained from the two raters in each test task were compared by using mean, standard deviation and *t*-test to determine the differences in the students' oral performance. The videotaped interactions from the two oral test tasks were additionally transcribed and analyzed following the Conversation Analysis (CA) methodology adopted by Seedhouse (2004). The transcribed data were closely examined with reference to different interactional features oriented to by the talk participants in the two oral test tasks.

Findings and discussion

Results from the two test tasks used in assessing the students' oral performance with the traditional rubric are displayed and discussed below.

Table 1: Score results from two types of assessment

Criterion	Interview interaction		Two-Party peer interaction		<i>t</i> -value	Sig.
	\bar{x}	S.D.	\bar{x}	S.D.		
Fluency	9.25	1.36	9.15	1.38	0.28	0.40
Vocabulary	8.90	1.20	9.20	0.71	-0.94	0.19
Appropriacy	9.75	1.14	10.00	0.88	-0.60	0.28
Comprehensibility	9.90	1.41	10.05	1.32	-0.34	0.37
Grammar	8.45	0.98	8.95	0.60	-1.94	0.04*
Total	46.25	5.28	47.35	4.24	-0.83	0.22

* Significant at 0.05 level

As shown in Table 1, the total scores obtained from both test tasks were not significantly different at the .05 level ($t = -0.83$, $p = 0.22$). It was suggested that both interview interaction and two-party peer interaction produce similar results overall when used in assessing the learners' conversational performance with the traditional linguistic-competency oriented rubric.

However, it should be noted that, although not statistically different, the students appeared to perform better in the peer interaction. The mean scores on discrete items such as vocabulary, appropriacy, and comprehensibility were higher than those obtained from the interview interaction. The higher mean scores on these aspects in peer interaction may be due to two facts peculiar to the test format. First, its looser structure provides more opportunities for students to use the full range of their English. Moreover, two-party peer interaction by definition is between participants of similar if not the same background, social position, etc., in this case, students from the same school and class. Therefore, prior knowledge of an interaction with the peers makes for easier communication and greater inter-subjectivity. The participants not only are more relaxed but also can perceive even predict their interlocutor's understanding or lack thereof from shared experiences, visual cues, etc.

Additionally, it was particularly interesting to note that, in terms of grammar, the students scored significantly better in their interaction with peers than with the native speaker. This indicated that the students were more conscious of their grammar when talking with peers. This may seem counter intuitive, but the students seemingly knew that their native speaker interviewer (their teacher) would be more tolerant and understand them regardless of their grammatical errors. On the other hand, with a peer interlocutor, knowing that they were being observed by a third party, the researcher, and probably fearing communication misunderstanding and losing face, they made an effort to produce grammatically accurate turns and often initiated self-repair on grammar.

As far as fluency is concerned, the students however scored better when interacting with the native speaker interviewer than with their peers, suggesting that the interviewer teacher, who played a dominant role in the interaction, appeared to be more effective in noticing and dealing with problematic, dispreferred turns. Through sequence opening and closing turns, topic/sequence extension, collaborative turn completion, repair initiation by asking for clarification, seeking confirmation, and supplying correct forms of words, gaps in and between turns seemed to be reduced, making the conversation sound more fluent. Therefore, in line with Davis (2009), it can be implied that that lower-proficiency speakers may benefit from working with an interlocutor who is more conversationally competent in the target language.

Regarding the interactional features of students' talk elicited by the two types of assessment, Table 2 demonstrates the main different interactional characteristics of the two test tasks.

Table 2: Interactional characteristics of the students' talks elicited from the two types of assessment tasks

Interview interaction	Two-party peer interaction
<i>Topic initiation/sequence opening</i>	
<ul style="list-style-type: none"> • No turns to establish a topic 	<ul style="list-style-type: none"> • Open question used to establish a topic • Balance between the two participants in topic initiation

Topic extension

- Limited reactive responses
- No topic extension
- Balance between the two participants in expanding a topic

Topic shift

- No self-initiated topics
- Balance between the two participants in topic shift

Turn size

- Unbalanced quantity of talk between the interviewer and interviewee
- Quantity of talk in the sequences balanced between the two participants

Turn construction unit

- Use gestures as well as verbal language to convey meaning
- Resorted to gestures to mime the target word

Explicit abandonment of a sequence

- No attempt to respond to an opening turn in an adjacency pair sequence at times of difficulty
- Supplying acknowledgement tokens (e.g. “yes”) or displaying agreement (e.g. “I think that too”)

Use of fillers

- Use gambits (e.g. er, hmm) to fill the pauses
- To gain time in order to keep the conversation going at times of difficulty

Overlap and latches

- Overlaps with minimal responses or acknowledgement tokens (e.g. “yes” or ((nods)))
- Overlaps with competitive interruptions

Repair

- Little collaborative completion
 - Asking for repetition
 - Collaborative completion
 - Asking for repetition
 - Asking for clarification
 - Seeking confirmation
-

It was revealed that in the interview interaction, since it was the interviewer who had rights to open and close the dialogue and to ask questions and introduce new topics, the students did not get to perform any of these actions. They were mostly prompted to answer the questions asked by the

interviewer (their teacher), and were not given much opportunity to negotiate for meaning as the teacher played a dominant role in turn initiation and allocation. The students in two-party peer interaction, on the other hand, made a collaborative effort to construct turns and maintain ongoing sequences. In line with Kormos (1999), through such interactional features as sequence openings and closings, introducing new topics and topic shift, the students were apparently able to exercise their most basic rights in conversation with their peers.

As shown in the table, the students engaged in the interview interaction task had no opportunity to establish or initiate the topic of the talk. It was the interviewer who always proffered the topic via open and closed questions to which the students provided responses by recycling and topicalizing the information provided by the interviewee. In contrast, in two-party interaction, the students could balance both roles: topic initiators and responders and participate in a more reciprocal exchange.

In addition, the students engaged in interview interaction also had little chance to reformulate trouble turns since most of the turns were repaired by the interviewer. When interacting with peers, they showed an attempt to repair the trouble turns or shift the topic if the attempt failed.

In the interview assessment, since the abilities to initiate or extend the topic were controlled by the interviewer, the students' turns seemed to be smaller and less complex than the interviewer's. In the interview interaction, the teacher obviously emerged as the more dominating speaker as he or she spoke more, took longer turns, made all the initiation moves, and was the only speaker to expand on his or her ideas. In contrast, the quantity of talk was balanced between the two participants in peer tasks. The rights and duties of the interactants were apparently equally distributed in the speakers' turns.

The students in interview interaction also employed more gesture-only turns. In addition, when confronted with difficult questions asked by the teacher, the students tried to respond to the question first with gestures and later with words even though they mostly failed to produce a complete turn, ending up with a smile and a pause. Similarly, struggling with certain words, they also resorted to gestures to mime the target word in the peer interaction.

Despite the failure to finish an ongoing sequence, when interacting with peers, the students were found to resort to strategies such as using an acknowledgment token, e.g., "yes," a filler, e.g., "hmm" or an agreement token, e.g. "I think that too" to gain time in order to keep the conversation going at time of difficulties.

The talks elicited from the two test tasks also differ as far as overlaps are concerned. In the case of two-party peer interaction, the frequent overlaps did not result in a topic shift but an extension of the prior topic or a support for the speaker who initiated the topic. On the other hand, in the interview interaction, the students proffered only minimal responses such as "yes" or a nod.

Another saliently different feature of the interaction elicited from the two test tasks lies in repair organization. The data transcribed showed that the important type of repair organization occurred in both test tasks was repair for meaning and fluency. Linguistic errors were often ignored in both test tasks if they did not obscure the meaning. The repair strategies used varied depending on the tasks. Collaborative completion was not found at all in the interview task while a wider range of repair strategies were resorted to in the two-party peer interaction task such as collaborative completion, asking for repetition, asking for clarification and seeking confirmation.

Based on the findings, it can be suggested that either interview or peer interaction test task can be used to assess students' linguistic competency in oral communication since they do not produce

significantly different results when used with the traditional rubric. However, as far as conversational competencies are concerned, the interview task may be of limited value since it does not allow the students to produce the interactional features sufficiently characterizing their conversational competence, which is a crucial component of EFL oral proficiency.

Being an unequal social encounter, the interview task does not create a situation in which natural conversations can occur. The two-party peer interaction, on the other hand, provides the candidates more opportunities to perform relevant interactional features at talk particularly those related to organizing adjacency pair sequences, openings, closings, initiating, extending, and shifting topics, as well as dealing with problems in talk.

Conclusion

This study examined whether the two types of assessment produce different results when used in assessing Thai EFL students' oral performance with a traditional rubric. On the basis of the comparative analysis of the talk elicited through the two test tasks, i.e., interview interaction and two-party peer interaction, it can be concluded that using the traditional rubric, the total scores obtained from the two test tasks were not different even though the students in the peer interaction task performed better in most of the discrete items. These may be due to the fact that the students were provided more opportunities to formulate their own turns and felt more relaxed when conversing with their peer.

The results also suggest that in terms of the interactional features elicited the interview interaction may not be appropriate for assessing students' conversational competencies. Due to the unequal distribution of power, candidates in the interview interaction rarely have the chance to initiate a new topic and have no right to initiate the opening or closing of the conversation, which are fundamental to conversational ability. In contrast, peer interaction is recommended as a better means of assessing candidates' conversational competence.

Recommendations

Several directions and recommendations for further studies can be outlined below. In order to determine the reliability of the different test tasks used for measuring students' oral proficiency, there is a need for more controlled studies. In such studies, the 1-5 rating scale should be made clear to the raters by specifying correspondent speech features. Additionally, even though unstructured peer interview tasks may be able to elicit more features of natural conversation, future studies should select familiar topics which interest the participants and employ such tasks as non-scripted role-play to elicit their talk. In role-play activities, participants are given more freedom to initiate the conversation with their own turns without having to be concerned with the list of prescribed questions and are constrained only by specific roles and speech situations, more approximating natural conversation.

Furthermore, rather than focusing primarily on linguistic competencies, future studies which desire to assess learners' conversational competencies with more accuracy should also consider using a

rubric which includes such interactional features as reported. In addition, it would be worthwhile to examine how learners of different levels of English proficiency perform in the task using such a rubric.

Thus far, this study has not only shed light on the design of more appropriate speaking tasks and rubrics to assess learners' conversational competencies but also helped raise teachers' awareness of the interactional features common to natural conversations, which really need to be taken into account if we want to make the prevalent oral proficiency test more valid and conversation teaching and learning more effective (Fujii, 2012).

References

- Bachman, L. (1990). *Fundamental consideration in language testing*. Oxford: Oxford University Press.
- Davis, L., (2009). The influence of interlocutor proficiency in a pair oral assessment. *Language Testing*, 26, 367-396.
- Ducasse, A.M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing*, 26(3), 423-443.
- Egyud, G., & Glover, P. (2001). Oral testing in pairs – A secondary school perspective. *ELT Journal*, 55(1), 70-76.
- Fujii, Y. (2012). Raising awareness of interactional practices in L2 conversations: Insights from Conversation Analysis. *International Journal of Language Studies* 6(3), 99-126.
- Galaczi, E.D. (2008). Peer-peer interaction in a speaking test: The case of the first certificate in English examination. *Language Assessment Quarterly*, 5(2), 89-119.
- Hughes, A. (2003). *Testing for language teacher*. Cambridge: Cambridge University Press.
- Iwashita, N. (1996). The validity of the paired interview format in oral performance assessment. *Melbourne Paper in Language Testing*, 5, 1-65.
- Kormos, J. (1999). Simulating conversations in oral proficiency assessment: A conversation analysis of roleplays and non-scripted interview in language exams. *Language Testing*, 16(2), 163-188.
- Lazaraton, A. (1992). The structural organization of a language interview: A conversation analytic perspective. *System*, 20, 373-86.
- O'Loughli, K.J. (2001). *Study in language testing 13*. Cambridge: Cambridge University Press.
- Seedhouse, P. (2004). *The interactional architecture of the language classroom: A Conversation Analysis perspective*. United Kingdom: Blackwell Publishing.
- Underhill, N., (1998). *Testing spoken language*. Cambridge: Cambridge University Press.
- Van Lier, L. (1989). Reeling, writhing, drawling, stretching and fainting in coils; oral proficiency interviews as conversation. *TESOL Quarterly*, 23, 489-508.

Weir, C.J., (1993). *Understanding and developing language tests*. United Kingdom: Prentice Hall

Young, R. (1995). Conversational styles in language proficiency interviews. *Language Learning*, 45, 3-42.

Appendix 1

Scoring criteria and descriptors

Fluency	Vocabulary	Appropriacy	Comprehensibility	Grammar
Students will be able to speak fluently with only occasional hesitation. Speech may be slightly slower than that of a native speaker.	Students will be able to use vocabulary broad enough to allow the candidate to express most ideas well.	Students will be able to demonstrate good awareness of social conventions and have some abilities to respond to unpredictable turns in conversation, though may sound unnatural.	Students will be able to understand most speech or may require repetition of details.	Students will be able to generally demonstrate control of varieties of structures with only minor errors.

PAPER 2

A Comparative Study of Features of Talk Elicited by Two-Party Peer Interaction and
Oral Interview Tests Implemented with Thai EFL Learners

การศึกษาเชิงเปรียบเทียบลักษณะทางภาษาจากแบบทดสอบทักษะการสนทนา
กับเพื่อน และการสนทนาแบบสัมภาษณ์ ของนักเรียนไทย
ที่เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ

**A Comparative Study of Features of Talk Elicited by Two-Party Peer
Interaction and Oral Interview Tests Implemented with
Thai EFL Learners**

รัชต์วารณ อุศมา¹, ดร.เข็มทอง สินวงศ์สุวรรณ²

¹นักศึกษามหาบัณฑิต หลักสูตรการสอนภาษาอังกฤษเป็นภาษานานาชาติ คณะศิลปศาสตร์
มหาวิทยาลัยสงขลานครินทร์, ²อาจารย์ คณะศิลปศาสตร์ มหาวิทยาลัยสงขลานครินทร์

Abstract

Although producing similar results with a traditional scoring rubric, oral interviews and peer interaction reportedly differ in their ability to assess candidates' ability to participate in normal conversation. Through close analyses of talk elicited from Thai high school students' participating, this paper aims at delineating interactional features contributing to the difference between the two test tasks in assessing conversation abilities. The participants of the study were ten English-program students at Pimanpittayasan School, Satun who were engaged in 5-minute interview and two-party peer interactions. The conversations obtained were videotaped, transcribed and analyzed according to Conversation Analysis (CA) principles. The findings revealed that the two test tasks induced different interactional features some of which were related to learners' conversational problems. The main ones illustrated in the paper are sequence opening, extending and closing, sequence abandonment, turn size, gesture-only turns, overlap and repair initiation. It was suggested that based on the different interactional features elicited, the interview interaction may not be appropriate for assessing students' conversational competencies.

Keywords: *Conversation competency assessment; interview interaction; two-party peer interaction; Conversation Analysis (CA); Thai EFL learners*

บทคัดย่อ

การประเมินทักษะการพูดโดยใช้วิธีการสัมภาษณ์ กับวิธีการจับคู่สนทนาระหว่างเพื่อนนักเรียนไม่แตกต่างกันอย่างมีนัยสำคัญ เมื่อใช้ทดสอบความสามารถในการสนทนาภาษาอังกฤษ โดยใช้ประเด็นการประเมินรูปแบบเก่า แต่เมื่อวิเคราะห์ลักษณะการโต้ตอบของผู้พูดในบทสนทนา ที่ปรากฏจากทั้งสองรูปแบบการประเมินพบว่ามีความแตกต่างกัน บทความฉบับนี้จึงมีวัตถุประสงค์ เพื่อชี้ให้เห็นความแตกต่างดังกล่าว โดยกลุ่มตัวอย่างที่ใช้เก็บข้อมูลเพื่อการวิจัยในครั้งนี้ คือนักเรียนที่เรียนหลักสูตรการจัดการเรียนการสอนเป็นภาษาอังกฤษ ชั้นมัธยมศึกษาปีที่ 3 โรงเรียนพินานพิทยาสรรค์ จังหวัดสตูล จำนวน 10 คน นักเรียนกลุ่มดังกล่าวเข้ารับการประเมินทักษะการพูดทั้งสองรูปแบบตามหัวข้อที่ครูกำหนดให้ โดยในแบบแรกนักเรียนถูกสัมภาษณ์โดยครูผู้สอนชาวต่างชาติ และแบบที่สองนักเรียนคุยโต้ตอบกับเพื่อนในหัวข้อเดียวกันกับแบบสัมภาษณ์ โดยใช้เวลาในการทดสอบอย่างละ 5 นาที โดยการทดสอบทั้งสองรูปแบบถูกบันทึกเทปเสียงบทสนทนาเพื่อการวิเคราะห์ข้อมูลตามหลัก Conversation Analysis (CA) ผลการวิจัยพบว่า ลักษณะการโต้ตอบในบทสนทนาสองรูปแบบมีความแตกต่างกัน ในประเด็นดังต่อไปนี้ sequence opening, extending and closing, sequence abandonment, turn size, gesture-only turns, overlap และ repair initiation ซึ่งความแตกต่างดังกล่าวชี้ให้เห็นว่าการประเมินทักษะการพูดสนทนาแบบวิธีการสัมภาษณ์ อาจไม่เหมาะที่จะนำมาใช้ทดสอบความสามารถในการสนทนาที่แท้จริงของผู้เรียน

คำสำคัญ: การประเมินความสามารถการสนทนา, การโต้ตอบแบบสัมภาษณ์, การโต้ตอบแบบจับคู่กับเพื่อน, การวิเคราะห์บทสนทนา, นักเรียนหลักสูตรภาษาต่างประเทศ

Introduction

As an essential part of any English language teaching curriculum, speaking skill is an important object of assessment. The purposes of a speaking assessment typically are to measure language proficiency; to assess achievement of the objective of a course; to diagnose learners' strengths and weaknesses; to identify what they do and do not know; and to assist in the placement of learners in a teaching program (Hughes, 2003). When assessing conversing ability, it is important to design a test that allows candidates to demonstrate their ability to use language in ways which are characteristic of their interactional competence. To test whether learners can converse, it is therefore necessary to get them to take part directly in speaking activities.

In an EFL classroom, there are several types of speaking test tasks that are commonly used to assess learners' speaking ability. An interview is one of the most popular means of testing speaking

skills (Underhill, 1998; Weir, 1993). It is a direct, face-to-face exchange between a learner and interviewer in which a learner's performance is evaluated. There is sometimes an assessor present who does not take part in the spoken interaction but listens, watches, and evaluates the abilities of the learner (Weir, 1993).

While widely practiced, there has been mounting criticism against the use of interviews in speaking assessment. Many researchers have come to agree that the oral exchange that occurs between an interviewer and a test taker does not reflect or even closely replicate natural or real-life conversation (e.g. Bachman, 1990; Lazaraton, 1992; Van Lier, 1989; Young, 1995). Interview tests often result in a *test discourse* or an *institutional talk* that neither represents normal conversation nor provides candidates with the opportunity to show their ability to participate in interaction other than as an *interviewee* responding to questions.

Consequently, since the late 1980s, as one of the most common practices in classroom activities, pair or group tasks have increasingly been used to assess speaking ability (Egyud & Glover, 2001). From a pragmatic perspective, pair or group assessment is typically also more time- and cost-efficient as learners are tested together, and raters assess two or more candidates simultaneously. When assessing students' oral performance in paired or group tasks, raters are oriented towards interactional or conversational management, recognizing the manner in which turns are verbally and nonverbally organized as an important feature contributing to successful interaction (Ducasse & Brown, 2009). It was argued that the paired test format makes possible the assessment of several salient features of talk and provides the potential for a wider range of interactional moves apparently reflecting genuine features of ordinary conversation.

Despite the argument that the interview was not as appropriate as paired peer interaction for measuring conversation skills in oral tests, it was too early to claim that either the interview interaction or the two-party peer interaction can better assess Thai EFL learners' oral performance. Given only few studies directly comparing the results of both oral proficiency test tasks, in the early phase of our research on Thai EFL learners, we attempted to determine whether the two test types produce different

results when assessing the learners’ oral performance with the traditional rubric oriented to features such as fluency, vocabulary, appropriateness, comprehensibility and grammar. The results revealed that both interview and two-party peer interactions produced similar score results when used in assessing the overall oral English performance of the learners with the traditional linguistic-oriented rubric.

However, close analysis of the recorded talks elicited from the Thai EFL learners investigated apparently revealed different interactional features emerging in the two-party peer interaction and oral interview tests. In this paper, we, therefore, examine those distinct features shaped by the two test tasks in detail.

Background

Traditionally, in an EFL classroom, oral tests, or tests to measure speaking ability, were focused largely on linguistic proficiency. However, recently, there has been an increase in the inclusion of conversation skills and strategies, as well as features of talk.

Clark (1979, as cited in O’Loughlin, 2001) provided the basis for distinguishing three types of speaking tests, namely, indirect, semi-direct and direct tests. *Indirect* tests generally refer to those tests which do not require the test taker to speak in language testing. Direct speaking tests, on the other hand, are those tests in which the test taker is asked to engage in a face-to-face communicative exchange with one or more interlocutors. The term *semi-direct* is additionally employed to describe those tests which elicit active speech from the test taker by means of tape recordings, printed test booklets, or other “*non-human*” elicitation procedures, rather than through face-to-face conversation with a live interlocutor.

Of the three speaking test types, i.e., direct, semi-direct, and indirect tests of oral proficiency, direct tests are generally the most valid procedures for measuring global speaking proficiency because of the close relationship between the test context and “*real life*” communication or face-to-face interaction. Nowadays, in EFL classrooms, direct speaking tests, therefore, seem to be preferred in assessing students’ oral performance.

While the interview test format is often chosen in order to assess the overall oral proficiency of candidates, peer interaction in the form of role-play is a popular choice in most Thai EFL classrooms. Each test format in fact has its strengths and weaknesses. Interview tests are widely practiced since they are not only easy to conduct but also able to elicit the language that displays what individual speakers

can do with what they have learnt, and how effectively they can communicate with native or near-native speakers (Peace Crops, 2005). However, according to many scholars (e.g. Bachman, 1990; Lazaraton, 1992; Van Lier, 1989; Young, 1995), the interview test does not reflect or even closely replicate natural or real-life conversation. It often results in a test discourse or an institutional talk that neither represents normal conversation nor provides candidates with the opportunity to show their ability to participate in interaction other than as an interviewee responding to questions. Consequently, pair or group tasks have increasingly been used to assess speaking ability (Egyud & Glover, 2001). In classroom contexts, they are typically more time- and cost-efficient as learners are being tested and assessed simultaneously.

Research Procedures

The participants in this study were ten high-proficiency students sampled from the population of 32 students in the M.3 English Program at Pimanpittayasan School, Satun Province. The students attended a Listening and Speaking course in the first semester of the academic year 2011 and took the oral tests which were part of their end-of-the year examination. Only high proficiency students were chosen for the investigation because they were believed to produce more talk to be assessed when working with partners (Iwashita, 1996).

As part of an achievement test to assess their oral performance during the course, the learners were first engaged in an oral interview interaction and at a later date in a two-party peer interaction on the same topic. Each interaction test was conducted three times over the course of three months. Prior to the interview interaction, the students were allowed to do research on the topic assigned by the teacher. The interviewer started the questions by discussing everyday topics and subsequently inquiring about the information the students had prepared. Each interview took approximately 5 minutes. In the two-party interaction assessment, on the other hand, the student peers were matched by the teacher. The interaction was assessed on the same criteria as the interview interaction. And just as in the interview, the students were asked to discuss the prepared topic for about 5 minutes. Both of the oral test tasks were video-recorded and transcribed following the Conversation Analysis (CA) convention for subsequent close analysis.

Findings and Discussion

Although both interview interaction and two-party peer interaction produce similar results overall when used for assessing the learners' conversational performance with the traditional linguistic-competency oriented rubric, close analysis of the talks elicited from the two test tasks indeed unveiled different interactional features some of which were related to learners' conversational problems. These features are illustrated and discussed in detail below.

Close comparative analysis of the conversations obtained from interview and peer interaction revealed differences in the following interactional features: sequence opening, extending and closing, sequence abandonment, turn size, gesture-only turns, overlap, and repair initiation.

Sequence opening, extension, and closing

In the interview interaction, since it was the interviewer who had rights to open, extend and close the dialogue and to ask questions and introduce new topics, the students did not get to perform any of these actions. As seen in Excerpt (1), the student was mostly prompted to answer the questions asked by the interviewer (her teacher), apparently not acknowledging the role as a conversation partner but a question-responder being assessed.

(1)

- | | | |
|----|-----|---|
| 1 | T: | <u>what</u> do you like about Thailand, |
| 2 | FS: | e:r I like food (.) Thai food (0.1) because it's delicious |
| 3 | T: | what is your favorite Thai food, |
| 4→ | FS: | em tom yum kung |
| 5→ | T: | what is the ingredient, do you know? |
| 6→ | FS: | no ((shaking her head)) |
| 7→ | T: | = tom yum is spicy soup an::d kung i:s what is kung in English e:::r? |
| 8 | FS: | I'm forget ((laugh)) |

- 8 T: praw:n it's ok,
 9 FS: (.)
 10 → T: what else do you like in Thai food
 11 FS ((confused face))

In this excerpt, the teacher initiated the talk through the opening question asking about the student's likes about Thailand. The student, in turn, responded with “food” and the reasoning increment “Thai food is delicious”. Expanding the question-answer sequence, the teacher further asked about her favorite food and the ingredients used to make it. The questions were responded to with a single phrase in line 4 and a dispreferred response in line 6. Having difficulty with the teacher's question at line 5, the student did not try to fix the problem but yielded the turn, letting the teacher continue with hers. Through the extended turn in line 7, the teacher apparently was hinting at the answer and herself having a problem finding the English word for the Thai one, “kung”. Receiving no help from the student, she completed the repair herself, resulting in self-initiated self-repair, and continued with a new question in line 10 given no further uptake from her student. Accordingly, just as in any typical interview interaction the interviewer or the teacher ended up playing a dominant role, doing most of the talk.

The students in two-party peer interaction, on the other hand, were on equal footing when taking turns. They opened, extended, and closed the sequences all by themselves. As seen in (2), the speakers reciprocally conducted the greetings, lines 1-2, and initiated the question-answering sequence, lines 3 and 8. No dispreferred responses were observed; the uptakes were promptly provided, as seen through the latching turns in lines 2 and 5. They made a collaborative effort to maintain the ongoing sequence till it came to a consensual close despite apparent problems, as seen in lines 11-15. In line with Kormos (1999), through these interactional features, the students were obviously able to exercise their most basic rights in conversation, and their rights and duties were apparently equally distributed in their turns.

(2)

- 1 → So: er::m good- good afternoon Suprawee =
- 2 → Pim: = good afternoon Soraida
- 3 → So: erm: what do you like in:: in Thailand
- 4 Pim: I like actor! =
- 5 → So: = actor. wow:: wo:o
- 6 Pim: new::s girl ((clap her hands))
- 7 So: wo::w
- 8 → Pim: and you?
- 9 So: erm: I like singer
- 10 Pim: why!
- 11 → So: because Thai- because Thai singer is (.)
- 12 I think- somebody is perfect, somebody is °not°
- 13 Pim: °yeah° ((nods))
- 14 So: yeah
- 15 → Pim: I think that too
- 16 So: yes:: erm:: what do you like- where do you like
- 17 erm: *mai chai* ((no it's not)) erm:: what do
- 18 you like Tourism, in Thailand,

Sequence abandonment

As previously seen in (1), while interacting with their teacher in the interview interaction, the students mostly proffered minimal verbal and non-verbal responses, e.g. “yes”, “no”, a nod, or simply waiting for the teacher’s move. Taking the dominant role, the teacher completely controlled the direction of the sequence. As seen below in excerpt (3), in line 6, instead of accepting the student’s response in line 5 and asking her to elaborate on it, the teacher chose to abandon it for a new question-answer sequence, starting at line 6.

(3)

- 1 T: as a local person living in Satun which places
 2 would you like to take a tourist to visit and why.
 3 FS: Tarutao
 4 T: why.
 5 → FS: near em:: kho lipe
 6 → T: what part is near th::e er:: kho tarutao
 7 FS: (0.6)
 8 T: do you know? which place in satun
 9 FS: °langu°
 10 T: langu? have you been in langu before?

Illustrated in excerpt (4), no such abandonment was found.

(4)

- 4 Toon: er:: I –I will ask you (.) as you are a er: local people
 5 which place you er::: you would like to- to take tourist to visit
 6 Amp: em::: I think the Phupa cave is the best to visit
 7 Toon: why? ((smile))
 8 Amp: because (.) it's (.) the famous in Thailand (.) it's a biggest
 9 cave in Thailand (.) have many things to [learn
 10 Toon: [yeah] em:: (0.3)
 11 I would like to go there

Turn size

Concerning turn size, as previously seen in (1), since the abilities to initiate, extend or close a sequence were controlled by the interviewer, the students' turns seemed to be smaller and less complex

than the interviewer’s. In the interview interaction, the teacher, who is more proficient in the target language, obviously emerged as the more dominating speaker as she spoke more, took longer turns, made all the initiation moves, and was the only speaker to expand on her ideas. In contrast, the quantity of talk was balanced between the two participants in the peer tasks, as in (4).

Gesture-only turns

The students in the interview interaction also employed more gesture-only turns. Particularly, when they were confronted with difficult questions, they tried to respond to the question first with gestures and later with words, often through the help of the conversation partner. The students, however, often failed to produce a complete turn, ending up with yet another gesture-only turn, as illustrated in (5).

(5)

- | | | |
|-----|-----|--|
| 1 | T: | as a local person living in Satun which places would |
| 2 | | you take a tourist to visit and why. |
| 3 | FS: | Wangsaithong er:: waterfall |
| 4 | T: | where is wangsaithong |
| 5 | FS: | er:: is in La-ngu (.) |
| 6→ | T: | what can you see the:: |
| 7→ | FS: | er:: em:: |
| 8→ | | (0.3) |
| 9→ | T: | can you describe it wangsaithong. |
| 10 | | (.) |
| 11 | FS: | I don’t °understand° |
| 12→ | T: | what dose wangsaithong look like |
| 13→ | FS: | er:: ((extending her hand out and smile)) |
| 14 | T: | river? |

- 15 FS: yes.
- 16 T: it has the river in that waterfall ((laugh))
- 17 FS: ((laugh)) yes yes ((nods))
- 18 T: ((laugh)) are you sure
- 19→ FS: ((smile and nods))
- 20 T: ok? wangsaihong how many level does it have.

As shown in Excerpt (5), through the dispreferred responses and the micropauses in lines 7-8, the student apparently did not understand the teacher’s question in line (6), leading the latter to rephrase the question in line 9, which yet failed to elicit a preferred response. Eventually at line 13, she responded to the teacher’s reformulated question in line 12 with gesture, miming the waterfall. A gesture-only turn was also produced in line 19 as a response to the teacher’s confirmation-seeking question.

Overlap

The talks elicited from the two test tasks also differ as far as overlaps are concerned. As seen in Excerpt (1), lines 4 and 6, and Excerpt (3), the students mostly proffered minimal responses with no overlap. However, in two-party peer interaction, frequent overlaps are observable showing alignment between speakers. Similar to lines 9 and 10, in Excerpt (4), in Excerpt (6), lines 24 and 25, Chon overlapped Ta to offer her agreement with the idea of HIV being a negative consequence of sex tourism. Such an action was not found at all in the interview interaction.

(6)

- 18 Chon: erm:: the business of polity to entertain for people who are holiday,
19 come to other country or in country twelve
20 that play the places ((extending her hands out))
21 what do you think about (0.1) sex tourism.
22 Ta: erm: bad very bad

- 23 Chon: why,
- 24→ Ta: will be have er:: HIV [in country
- 25→ Chon: [yes very bad erm:: do you think tourism
- 26 help people understand each other

Repair initiation

Another different salient feature of the interaction elicited from the two test tasks lies in repair organization. The data transcribed showed that the important type of repair organization occurred in both test tasks was repair for meaning. Repair initiation takes different forms in the two tasks. Collaborative completion and seeking confirmation were not found at all in the students' turns in the interview task. In peer-interaction, the students resorted to a wider range of repair-initiation strategies, not only seeking confirmation but also asking for repetition and clarification.

(7) (Collaborative completion)

- 18 Ta: what you dislike in Thailand? [about Tahiland
- 19 Chon: [em::
- 20 I:: I dislike some some locate many garbage
- 21→ Thailand to- too garbage em:: [pla:: plastic?
- 22→ Ta: [plastic? Yes
- 23 Chon: yeah for for in river yeah

In Excerpt (7), an instance of collaborative completion was observable. At line 21, Chon's search for the word “plastic” is collaboratively completed by her conversation partner in line 22. Such cases were not found in the interview interaction with the teacher.

Likewise, confirmation seeking was also found in the students' talk only in peer interaction. Shown in excerpt (8), at line 30, So repeated “all rock is black” by putting emphasis and raising the intonation on the rock's color in order to seek the confirmation of Pim's talk in line 27.

(8) (Seeking confirmation)

- 24 Pim: em:: is a- I would like to (0.4) oh?
 25 kho hin ngam? kho hin ngam has er:: rock=
 26 So: =rock wow how about rock
 27→ Pim: the rock is old the rock is black- black rock that's not [sand
 28 So: [oh? em::
 29 Pim: =it has er:: rock
 30→ So: er:: all rock is black?
 31 Pim: =yes and you as you are a- a local people which place
 32 wo- would you like to take the tourist to visit

Besides collaborative completion and confirmation seeking of a turn, the students in peer-interaction also resorted to such repair-initiation strategies as asking for clarification and repetition. Seen in Excerpt (9), in line 39 Chon requested an example from Ta to clarify the answer given in line 38 regarding the food she likes, after which she started to list the examples, completing the other-initiated self-repair.

(9) (Asking for clarification)

- 36 Chon: food what [what what do you like food
 37 Ta: [what food do you like
 38→ many. many [food I like,
 39→ Chon: [example example
 40→ Ta: em:: I like I like em Som Tum, =
 41 Chon: = ahh:: Tomyum Kung ((pointing to her friend)) =
 40 Ta = yeah: Tomyum Kung

Asking for repetition also occurred when the students talked with their peers and could not hear or understand their utterances properly. As can be seen in Excerpt (10), Jing in line 30 asked Rose to

repeat the question asked in lines 27 and 29. Such an instance occurred less frequently in the interview interaction with the teacher, only 3 out of 30 excerpts.

(10) (Asking for repetition)

- 24 Rose: would you like to work in tourism.
- 25 Jing: no I wouldn't because it's work hard to me (0.1)
- 26 I'm lazy::
- 27→ Rose: why. why it's work hard (.) what- what work is it
- 28 Jing: ah:: sometimes it's different (difficult) to me
- 29→ Rose: what work
- 30→ Jing: again please,
- 31→ Rose: what work
- 32 Jing: work [another
- 33→ Rose: [fisher or:: seller
- 34→ Jing: seller seller

Conclusion

The different interactional features found in the talks elicited from the two types of assessment apparently suggested that the interview interaction may not be appropriate for assessing conversational competencies. Due to the unequal distribution of power and the imbalanced language ability, test candidates played a subordinate role in sequence organization, thus less talk being produced for the assessment and the interaction being more like a question-answering session. When facing problems in the interview, the students also resorted to fewer repair strategies. In contrast, the students in peer interaction produced more balanced talk and drew on a greater range of interactional strategies. Thus the latter should be a better means of assessing conversational competence especially of low-proficiency students.

References

- Bachman, L. (1990). *Fundamental consideration in language testing*. Oxford: Oxford University Press.
- Ducasse, A.M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing*, 26(3), 423-443.
- Egyud, G., & Glover, P. (2001). Oral testing in pairs – A secondary school perspective. *ELT Journal*, 55(1), 70-76.
- Hughes, A. (2003). *Testing for language teacher*. Cambridge: Cambridge University Press.
- Iwashita, N. (1996). The validity of the paired interview format in oral performance assessment. *Melbourne Paper in Language Testing*, 5, 1-65.
- Kormos, J. (1999). Simulating conversations in oral proficiency assessment: a conversation analysis of role plays and non-scripted interview in language exams. *Language Testing* 16(2), 163-188.
- Lazaraton, A. (1992). The structural organization of a language interview: a conversation analytic perspective. *System*, 20, 373-86.
- O'Loughli, K.J. (2001). *Study in language testing 13*. Cambridge: Cambridge University Press.
- The Peace Crops Information Collection and Exchange. (2005). *Language proficiency interview (Manual for testers)*. Washington, DC: ICE Press.
- Underhill, N. (1998). *Testing Spoken Language*. Cambridge: Cambridge University Press.
- Van Lier, L. (1989). Reeling, writhing, drawling, stretching and fainting in coils; oral proficiency interviews as conversation. *TESOL Quarterly*, 23, 489-508.
- Weir, C.J. (1993). *Understanding and Developing Language Tests*. United Kingdom: Prentice Hall.
- Young, R. (1995). Conversational styles in language proficiency interviews. *Language Learning*, 45, 3-42.

Appendix

Transcription convention adapted from Seedhouse (2004) and Schegloff (2007)

[Point of overlap onset
]	Point of overlap termination
=	(a) Turn continues below, at the next identical symbol (b) If inserted at the end of one speaker’s adjacent turn, indicates that there is no gap at all between the two turns (c) Indicates that there is no interval between adjacent utterances
(0.5)	Numbers in parentheses indicate silence, represented in tenths of a second; what is given here indicates 5 second of silence
(.)	Very short untimed pause; ordinarily less than 2 second
<u>word</u>	Speaker emphasis
-	A hyphen after a word or part of a word indicates a cut-off or self-interruption
?	Rising intonation, not necessarily a question
.	Low-rising intonation, or final, not necessarily the end of a sentence
()	A stretch of unclear or unintelligible speech
wo:rd	Colons show that speaker has stretched the preceding sound
owordo	Material between “degree signs” is quieter than the surrounding talk
((word))	Transcriber’s comments
[gibee]	In the case of inaccurate pronunciation of an English word, an approximation of the sound is given in square brackets
<i>ja</i> ((tr.: yes))	Non-English words are italicized and followed by an English translation in double parentheses
→	Mark features of special interest

VITAE

Name Ms. Ratchawan Ussama

Student ID 5311121040

Educational Attainment

Degree	Name of Institution	Year of Graduate
Bachelor of Business Administration in Tourism Management- Ecotourism	Walailak University	2006

List of Publications

Ussama, R., & Sinwongsuwat, K. (2013). A comparative study of features of talk elicited by two-party peer Interaction and oral interview tests implemented with Thai EFL learners. *Proceedings of the Hatyai Symposium 2013 "Research for Social Development of Thailand"*, Thailand: Hatyai University.