To the memory of my grandmother and
grandfather whom I have never met…

Factors Influencing Dyadic Interaction in Paired Speaking Tests: Proficiency and Familiarity

The Graduate School of Education
of
İhsan Doğramacı Bilkent University

by

Esma Kot

In Partial Fulfilment of the Requirements for the Degree of
Master of Arts
in
Teaching English as a Foreign Language

Ankara

November 2017

İHSAN DOĞRAMACI BILKENT UNIVERSITY

GRADUATE SCHOOL OF EDUCATION

Factors Influencing Dyadic Interaction in Paired Speaking Tests: Proficiency and Familiarity

Esma Kot

November 2017

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Asst. Prof. Dr. Deniz Ortaçtepe (Supervisor)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Prof. Dr. Julie Mathews-Aydınlı (Examining Committee Member)

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Teaching English as a Foreign Language.

Asst. Prof. Dr. Zeynep Sağtaş-Bilki (Examining Committee Member)

Approval of the Graduate School of Education

Prof. Dr. Alipaşa Ayas (Director)

ABSTRACT

FACTORS INFLUENCING DYADIC INTERACTION IN PAIRED SPEAKING TESTS: PROFICIENCY AND FAMILIARITY

Esma Kot

M.A in Teaching English as a Foreign Language

Supervisor: Asst. Prof. Dr. Deniz Ortaçtepe

November, 2017

This study investigated whether familiarity and proficiency factors play a role in EFL learners' use of interactional resources (i.e., turn-taking, topic management, repair and task management) and the emergence of interactional patterns (i.e., collaborative, parallel, asymmetric and blend) in paired speaking tests. The study was carried out with 100 EFL learners paired as low-low, high-high and low-high and with 36 EFL learners matched as unfamiliar and familiar in the oral proficiency exam at a state university in Turkey.

In order to place the participants for low-low, high-high and low-high groups, their scores in the proficiency exam which measured their reading, writing and listening skills as well as their vocabulary and grammar knowledge, and their scores in the oral proficiency exam were examined by the researcher. While 15 pairs were selected for the low-low and high-high groups separately, 20 pairs were selected for the low-high group. Then all 100 students (50 pairs) in the first cohort were asked whether their partners were their classmates or their friends in the exam and nine pairs out of 50 were detected as familiar. After that, nine unfamiliar pairs were

selected in order to compare them with the familiar ones. In total, 50 videos were listened to and transcribed by using the conventions suggested by Jefferson (2004). Forthwith, all transcriptions were analyzed in order to identify the interactional resources such as turn-taking, repair, topic management and task management employed by the test-takers during the test discourse. Following this process, the researcher drew upon the interactional resources in order to assign the interactional patterns such as collaborative, asymmetric, parallel and blend which took place during test-takers' interaction with each other.

The results indicated that pairing two different proficiency level students is disadvantageous for the high level test-takers in terms of topic management, task management and repair. In contrast, while the low levels are advantageous in terms of topic and task management in particular, they are disadvantageous in terms of turn-taking. What is more, while high-high pairs create a collaborative pattern which is the most favorable one, low-low pairs usually create a parallel pattern. On the other hand, low-high pairs usually generate an asymmetric pattern due to the dominance of the high levels. Furthermore, the findings suggested that pairing two unfamiliar peer interlocutors seem more advantageous for the test-takers because unfamiliar pairs usually generate a collaborative pattern whereas familiar pairs usually create an asymmetric pattern during the test discourse.

In light of these findings, this study provided insights into how test-takers should be matched in paired speaking tests for the test administrators.

Key words: peer interlocutor, proficiency, familiarity, pairing system, paired speaking tests, interaction, interactional resources, interactional patterns

ÖZET

EŞLİ KONUŞMA SINAVLARINDA ÇİFT TARAFLI ETKİLEŞİMİ ETKİLEYEN FAKTÖRLER: YETERLİLİK SEVİYESİ VE AŞİNALIK

Esma Kot

Yüksek Lisans, Yabancı Dil Olarak İngilizce Öğretimi

Tez Yöneticisi: Yrd. Doç. Dr. Deniz Ortaçtepe

Kasım, 2017

Bu çalışma aşinalık ve dil yeterliliği faktörlerinin İngilizceyi yabancı dil olarak öğrenen öğrencilerin, eşli konuşma sınavlarında kullandıkları etkileşimsel kaynaklar (yani; konuşma sırası, konu yönetimi, düzeltme yapma ve görev yönetimi) ve ortaya çıkan etkileşim modelleri (yani; işbirlikçi, parallel, asimetrik ve karma) üzerindeki rolünü incelemeyi amaçlamaktadır. Çalışma Türkiye'deki bir devlet üniversitesinde, yeterlilik seviyelerine göre sözlü konuşma sınavında düşük-düşük, yüksek-yüksek ve düşük-yüksek olarak eşleşmiş olan 100 öğrenci ve birbirine aşina ve yabancı olarak eşleşmiş 36 öğrenci ile gerçekleştirilmiştir.

Katılımcıları düşük-düşük, yüksek-yüksek ve düşük-yüksek gruplarına yerleştirmek için onların konuşma, yazma ve dinleme becerileri ile kelime ve dilbilgisi bilgilerini ölçen yeterlilik sınavındaki notları yanı sıra sözlü konuşma sınavında aldıkları notlar dikkate alınmıştır. Düşük-düşük ve yüksek-yüksek grupları için 15'er çift seçilirken, düşük-yüksek grubu için 20 çift seçilmiştir. Sonrasında toplam 100 (50 çift) öğrenciye konuşma sınavında partner oldukları kişiyle bir

yakınlığı olup olmadığı sorulmuştur ve 9 çiftin birbirlerini tanıdığı görülmüştür. Bu seçilmiş olan 9 aşina çiftle karşılaştırmak üzere birbirini tanımayan 9 çift daha seçilmiştir. Toplamda 50 video araştırmacı tarafından dinlenmiş ve Jefferson (2004) 'ın tavsiye etmiş olduğu semboller dikkate alınarak konuşmalar yazıya dökülmüştür. Daha sonra, yazıya dökülmüş olan konuşmalar etkileşimsel kaynakları belirlemek üzere analiz edilmiştir. Bunu takiben, araştırmacı etkileşim modellerini tanımlayabilmek adına bir önceki aşamada belirlenmiş olan etkileşimsel kaynaklardan yararlanmıştır.

Çalışmanın sonuçları farklı iki seviyedeki öğrenciyi eşleştirmenin yüksek seviyedeki öğrenciler için konu yönetimi, görev yönetimi ve düzeltme yapma kaynakları açısından dezavantajlı olduğunu göstermiştir. Bunun aksine, düşük seviyedeki öğrenciler için bu durum özellikle konu ve görev yönetimi açısından avantajlı bir duruma dönüşmektedir, fakat düşük seviyedeki öğrenciler de yüksek seviyede bir öğrenciyle eşleştiklerinde konuşma sırası açısından dezavantajlı durumdadırlar. Dahası, yüksek-yüksek olarak eşleşmiş öğrenciler en olumlu model olan işbirlikçi etkileşim modelini oluştururlarken, düşük-düşük olarak eşleşmiş öğrenciler paralel bir etkileşim modeli sergilemektedirler. Öte yandan, yüksek-düşük olarak eşleşmiş çiftler, yüksek seviyede olanların baskın olması sebebiyle, asimetrik bir etkileşim modeli oluştururlar. Ayrıca sonuçlar birbirini tanımayan iki öğrenciyi eşleştirmenin daha avantajlı olduğunu öne sürmektedir çünkü birbirini tanımayan öğrenciler sınav esnasında işbirliği içinde çalışırlarken, birbirine aşina olan çiftlerde konuşmacılardan birinin daha baskın olduğu ve bu sebeple asimetrik bir etkileşim modeli oluşturdukları gözlemlenmiştir.

Bu bulgular konuşma sınavı hazırlayanlar için eşli konuşma sınavlarındaki eşleştirme sistemi hakkında iç görü sağlamaktadır.

Anahtar sözcükler: eş konuşmacı, dil yeterliliği, aşinalık, eşleştirme sistemi, eşli konuşma sınavları, etkileşim, etkileşimsel kaynaklar, etkileşim modelleri

ACKNOWLEDGEMENTS

My one-year MA TEFL experience and writing this thesis involved both weal and woe. However, the woes turned into weal thanks to some individuals in my life. They were always there to encourage and support me. I would like to express my sincerest gratitude to those who accompanied me in this challenging process.

First and foremost, I would like to thank my thesis advisor Asst. Prof. Dr. Deniz Ortaçtepe for her excellent guidance and constructive feedback. Whenever I got confused, I knew that she would enlighten me with her wisdom. She was always there when I needed her. I feel privileged because she was my mentor. Without her support and guidance, it definitely would not be possible to complete this thesis.

I am also very grateful to Asst. Prof. Dr. Olcay Sert for sparing me his time and helping me determine this topic which was one of the hardest parts of this thesis. He was like a second advisor to me at the very beginning of this process. I would also like to express my gratitude to my committee members Prof. Dr Julie Mathews-Aydınlı and Asst. Prof. Dr. Zeynep Sağtaş-Bilki for their insightful comments and contributions to my study.

I also wish to express my gratitude to my institution, Bülent Ecevit University, Prof. Dr. Mahmut Özer, the President, and Prof. Dr. M. Haluk Güven, the Vice President, for giving me the permission to attend this distinguished master's program. I am also grateful to the Director of the School of Foreign Languages, Inst. Okşan Dağlı for giving me the permission to attend Bilkent MA TEFL and use the archives of the school to collect data.

I owe my deepest gratitude to Demet Kulaç Püren and Burcu Ak Şentürk who supported me with their guidance especially while applying to this program. I would like to extend my particular gratitude to Demet Kulaç Püren who livened my Mondays up in Ankara. It was great to meet her on some Mondays and to feel away from my stress for a while. I also owe many thanks to Gulçin Özgöz Gülenç, Bahar Bıyıklı Koç and Nihan Güngör Ün for their encouragement and support before I started this journey, in particular.

And my precious classmates, MA TEFLers! Special thanks to Şeyma Kökcü, Kamile Kandıralı, Kadir Özsoy, Güneş Tunç, Tuğba Bostancı and Nesrin Atak (the MA TEFLer of hearts) for their invaluable friendship and constant encouragement throughout this challenging year. There were times when we felt hopeless, but we always encouraged each other because we knew that these tough times could be overcome by cooperation. This wonderful team made this year meaningful and bearable. I am so glad to know all of you and I believe that our friendship will last forever with the help of our 'next' plans. I would also like to express my sincerest gratitude to Damla Koç and Hale Nur Peneklioğlu for their support. It was great to live in the same city with them for one year and to forget about my responsibilities when I met them.

I also owe many thanks to UDK which made me the person who I am today and to the people there. Since this club has contributed to me a lot in terms of the discipline and diligence, it deserves to be mentioned here.

I would like to express my heartfelt gratitude to my dear fiancé Turan Can Artunç for his endless love and never-ending faith in me from the very beginning to the end. I have felt much stronger since I met him. It may be because my generous hero shares his luck with me or because he reminds me who I am. I know that I can

overcome all difficulties as long as he holds my hand. I would also like to thank him for coming to the U.S. with me and for letting me experience the happiest moments of my life with his proposal in San Francisco.

Last but not least, my earnest gratitude goes to my beloved family; my father Erol Kot, my mother Döne Kot and my sister Esra Kot for their unconditional love and unquestioning support. They have always believed in me and helped me believe in myself. They have always supported me every time when I have set a new goal. I know that they are always there whenever I need them and I feel so lucky to have such a family! I am particularly indebted to my dearest sister, Esra Kot for being my best friend and life-long mentor. I always depend on her maturity and consideration. Even though we have been too far away from each other, I know that she is always with me.

Thank you all who are somehow in my life! You have made this thesis possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

Introduction

Designing speaking tests has always been a hotly debated issue in language testing. While discrete-point tests or techniques such as dictation and reading aloud were used for the purpose of assessing learners' speaking skills in the past, the rise of communicative language teaching in the 1970s led to an increase in performance tests which require the test-takers to have a conversation with an interlocutor. Even though individual interviews gained acceptance as a norm to test speaking skill at first, the emphasis on pair and group work in the classroom due to the communicative movement (Taylor & Wigglesworth, 2009) and the advent of interactional competence (Kramsch, 1986) have brought paired speaking tests into prominence.

Brooks (2009) claimed that paired speaking tests allow the test-takers to engage in more interaction and negotiate the meaning more in contrast to individual interviews. Still, the co-constructed nature of interaction in these tests brings about some arguments related to their implementation. Since the test-takers' performances are linked to each other, there is a possible interlocutor effect during the test discourse. Therefore, test-taker characteristics such as familiarity with partner, proficiency, gender and personality have been a concern in implementing paired speaking tests. Hence, this study will investigate whether interactant related factors (i.e., familiarity and proficiency) play a role in a) EFL learners' use of interactional resources and b) the emergence of interactional patterns in paired speaking tests.

Background of the Study

L2 oral proficiency can be defined as test-takers' ability to communicate with interlocutors, which aligns with Bachman's (1990) theoretical framework of communicative language ability. Upon the advent of the communicative movement, communicative competence became the focus of language teaching and offered some alternatives to the language testing field. This communicative approach to speaking tests led to the predominance of performance tests, particularly individual tests such as interviews (McNamara & Roever, 2006). Several techniques such as picture description, face-to-face interaction and oral proficiency interviews have been used for the purpose of assessing language learners' speaking skills (Birjandi, 2011). However, once interactional competence was proposed as an alternative theoretical framework to communicative competence (Sun, 2014), the established notions of communicative competence were challenged (Galaczi, 2013) and pair and group work activities gained importance in the language learning context (Taylor & Wigglesworth, 2009). Accordingly, paired speaking tests have been brought to forefront. In a paired speaking test, instead of being interviewed by an examiner who acts as an interlocutor, test-takers interact with each other in the presence of two examiners one of whom is an interlocutor and the other one is an assessor. This format possesses some advantages such as having the feature of leading to positive washback in the classroom (Messick, 1996) since it increases the use of pair and group work in language classrooms. In addition, test-takers may feel more comfortable since they share their anxiety (Saville, & Hargreaves, 1999). Additionally, paired speaking tests allow more interaction and negotiation of meaning during the test discourse (Sandlund, Sundqvist & Nyroos, 2016). Therefore,

interlocutors may have the chance to display their interactional competence more effectively in these tests.

In the field of SLA, interactional competence was first introduced by Kramsch (1986) who claimed that the main goal of developing students' proficiency in a foreign language should be to improve their interactional competence. It was later addressed by He and Young (1998) within Interactional Competence theory and they stated the importance of participating in 'interactive practices' so as to have interactional competence. (p.7) Young (2011) stated that "Interactional competence is not the knowledge or possession of an individual person, but is co-constructed by all participants in a discursive practice". (p.428) In other words, interactional competence is not the responsibility of an individual but it is jointly constructed. During the interaction both speakers need to contribute to the conversation and use interactional resources such as turn-taking, repair, and topic management, which will allow the speakers to generate a particular interactional pattern during the conversation.

Interactional patterns were first identified by Storch (2001, 2002) in his study on ESL writing pairs as; collaborative, dominant/passive, expert/novice, and dominant/dominant. Drawing upon Storch's (2001,2002) model, Dimitrova-Galaczi (2004) observed three interactional patterns during a paired speaking test and termed these patterns as collaborative, asymmetric, parallel and blend. The researchers suggested that in a collaborative pattern, interlocutors extend their own topics and are interested in their partners' topics as well. In other words, they contribute to the talk cooperatively. On the other hand, in a parallel pattern, test-takers develop their own topics, but they do not extend their partners' topics. Moreover, in an asymmetric pattern of interaction, there is an imbalance in the talk as the pairs in such a pattern

are observed as one dominant and one passive. Lastly, if the interlocutors generate two patterns together during their talk, it is referred to as a blend pattern (Galaczi, 2008). It was verified in Galaczi's (2008) study that test-takers who get involved in the collaborative pattern get the highest scores from assessors. What is more, to be able to generate the collaborative pattern, interlocutors need to know how language is used in talk-in-interaction and have a good grasp of interactional resources such as turn taking, repair, and sequence organizations that underlie all talk-in-interaction (Markee, 2008). The features of turn-taking management are turn length, speed and dominance (Dimitrova-Galaczi, 2013; Ducasse & Brown, 2009; Ducasse, 2010; Watanabe & Swain, 2007). More specifically, turn length means the utterance length or the number of turns a speaker takes and turn speed refers to how fast the two partners respond to each other. Furthermore, how interlocutors compete for the floor indicates turn domination (Ducasse & Brown, 2009). When it comes to repair strategies, studies have identified several repair strategies such as self-initiated self-repair, self-initiated other repair, other-initiated self-repair, other-initiated other repair, repetition, paraphrase, confirmation checks, clarification requests and comprehension checks (Drew, 1997; Schegloff et al., 1977; Schegloff, 2000). Repair is a term which helps speakers fix communication breakdowns. These strategies do not serve only the purpose of correction, but they are also used to deal with problems resulting from the mishearing and misunderstanding. On the other hand, task management refers to the extent to which interlocutors help each other to accomplish the task. In order to complete the task successfully, interlocutors should meet the requirements of the task. As for topic management, topic extension and topic shift are some of the concepts which constitute it. While in a conversation, interlocutors need to help each other to initiate and build upon a topic to maintain the

conversation. If the interlocutors do not extend the topic or resort to topic shift when necessary, it may impair the interaction between them. Thus, employing all those interactional resources such as repair, turn-taking and topic management is essential in order not to cause communication breakdowns in daily interactions.

Paired speaking tests may seem simple to implement; however, in an actual assessment context it can be extremely complex since test-takers' performance may be influenced by various tasks, examiners or interlocutors (Davis, 2009). From the mid-1990s, testers have been challenged by the issue that the spoken language displayed in a test is influenced during co-constructed interaction (He & Young, 1998 ; Jacoby & Ochs, 1995; McNamara, 2001) and these challenges have led to an increase in empirical studies of the nature of oral proficiency exams (Ducasse, 2009). Particularly, the effect of the other test-taker has been a big concern in paired speaking tests as both of the test-takers contribute to the interaction and their performances are linked to one another (Luoma, 2004; McNamara, 1997; Weir, 2005). As McNamara (1996) suggested "the age, sex, educational level, proficiency or native speaker status and personal qualities of the interlocutor relative to the same qualities in the candidate are all likely to be significant in influencing the candidate's performance". (p.86) Since the interaction is jointly constructed in paired speaking tests, the interactional resources employed by each interlocutor may influence the interaction that occurs during the test discourse. Moreover, the raters determine the scores looking at the interactional resources employed by test-takers and types of interactional patterns may have an impact on the rater scores (Galaczi, 2008). Therefore, peer interlocutor is an important source of variation which may either positively or negatively influence the discourse of the exam, and therefore a test-takers' performance and/or score (Csepes, 2002).

Statement of the Problem

Paired speaking tests not only better resemble natural conversation (Ducasse & Brown 2009), but also give test-takers the possibility of producing more varied patterns of conversation (Birjandi, 2011). Nevertheless, whether paired tests are more advantageous than traditional face-to-face interviews is a controversial issue, since the interaction is jointly constructed in these tests (Young & He, 1998). Furthermore, several researchers have suggested a possible interlocutor effect on test-takers' scores (Davis 2009; Galaczi 2008; Iwashita, McNamara & Elder, 2001) and on interaction (Lazaraton & Davis, 2008) during the test-taking discourse. Variables such as gender (e.g., Brown & McNamara, 2004; O' Loughlin, 2002; O' Sullivan, 2000), personality (e.g., Berry, 2004, 2007), familiarity with partner (e.g., Norton, 2005; O' Sullivan, 2002), and individual's proficiency (e.g., Davis, 2009; Dobao, 2012; Gan, 2010; Lazaraton & Davis, 2008; Nakatsuhara, 2006) have been found as factors which affect the test-takers' performances to some extent. Among these test-taker variables, especially proficiency, personality and interlocutor familiarity on peer-to-peer interaction appear to have received the most extensive attention (Van Moere, 2014). While the studies investigating familiarity have mostly explored its role on test-takers' linguistic performances, which have been measured by quantitative methods (e.g., O'Sullivan, 2002), few studies have been conducted on how actual interaction is affected by test-takers' proficiency levels (e.g., Galaczi, 2014; Gan, 2010). More specifically, to the best knowledge of the researcher, no studies have been carried out to explore whether pairing test-takers randomly without considering their proficiency levels or familiarity with each other would play a role in their use of interactional resources such as repair, turn-taking, topic management

and task management and the emergence of interactional patterns during the test discourse.

At the local level, in a state university in Turkey, for oral exams, students are paired randomly, regardless of their proficiency levels or familiarity with each other. Much as students take the placement exam at the beginning of the first and second semester, this exam does not measure their interactional competence. Therefore, there is the possibility to pair a low level student with a high level student in speaking tests. Galaczi (2014) claims that low level students have difficulty in engaging in interactions with their interlocutors. This means there is the risk that low level students who are paired with high levels may not be able to use the interactional resources and pairing system may affect the interaction between test-takers. On the other hand, while some students are paired with their classmates, others are matched with a stranger, which may turn out to be an advantage or a disadvantage in the test discourse. The aforementioned problems may affect the interaction that takes place in paired speaking tests. Once the interaction is affected by any of these factors, interlocutors may decrease each other's performances or assessors may distort the scores, which will destroy the reliability of the test results in the end.

Research questions

The present study aims to investigate whether interactant related factors (i.e., familiarity and proficiency) play a role in a) EFL learners' use of interactional resources and b) the emergence of interactional patterns in paired speaking tests. To achieve this aim, this study approaches interaction from two angles; first, interactional resources such as repair strategies, turn-taking, topic management and task management, and second, interactional patterns such as asymmetric, parallel, and collaborative and blend. In this sense, the research questions are;

1) How do the use of interactional resources and the emergence of interactional patterns vary in paired speaking tests with EFL test-takers

a) who have the same vs. different proficiency levels?

b) who are familiar vs. unfamiliar with each other?

Significance of the Study

Inasmuch as the interaction is co-constructed in dyadic tests, it can be influenced by factors such as proficiency and familiarity in a positive or negative manner. In this regard, this study may contribute to the field by indicating various interactional resources employed by low and high level test-takers and it may shed light on whether pairing EFL test-takers regardless of their proficiency levels and their familiarity with each other plays a role in the use of these resources and interactional patterns occurring during the test discourse. Hence, it may strengthen the theoretical basis of interactional competence and its components such as interactional patterns and interactional resources in the literature.

It may also pay dividends to the local institution, especially to the test administrators, in terms of how they should implement paired speaking tests. It may provide insights into whether the test administrators need to take into account the proficiency and familiarity factors while pairing the students so that test-takers can take the opportunity to co-construct a meaningful interaction in the test discourse. In addition to offering implications at the testing level, it may also offer some suggestions to language teachers at the teaching level. Being aware of the interactional resources, teachers can teach those resources in the classroom setting and address the communication problems of students more effectively. Furthermore, curriculum and material development units may benefit from this study to integrate these resources into the syllabus.

Conclusion

In this chapter, a brief introduction to the literature on paired speaking tests together with their potential advantages and disadvantages and interactional competence along with interactional resources and interactional patterns have been provided. Furthermore, the background of the study, the statement of the problem, research questions, and the significance of the study has been covered. The next chapter will present the relevant literature on oral proficiency interviews, interactional competence and its related concepts, and factors which influence the test discourse in paired speaking tests.

CHAPTER II: LITERATURE REVIEW

Introduction

The aim of this chapter is to review the current literature in relation to this study which investigates the role of the proficiency and familiarity factors on interaction in paired speaking tests. The literature will be discussed in three sections. In the first section, starting with a brief introduction to language testing, testing speaking and types of oral interviews are presented. After explaining the features of each oral interview format, their advantages and disadvantages are covered separately with an extensive focus on paired speaking tests. In the second section, the literature on interactional competence and related concepts such as interactional resources and interactional patterns are presented. In the last section, after mentioning the factors which affect the test discourse, a summary of related studies on factors such as task effect and interlocutor effect are presented separately with a particular focus on test-taker proficiency and familiarity.

Language Testing

The language testing field has undergone some changes up to the present time. For example, between the 1960s and the 1970s language testers designed tests which assessed learners' ability "as consisting of skills (listening, speaking, reading and writing) and components (e.g. grammar, vocabulary, pronunciation) and an approach to test design that focused on testing isolated 'discrete points' of language" (Bachman, 2000, pp. 2-3). During these years, the focus was on testing single units (Oller, 1979). That is to say, the tests designed to gauge learners' language

knowledge tested only one item of the language at a time. The 1980s witnessed the impact of communicative competence, which brought on the start of the shift from discrete-point tests to performance-based tests with the idea of using communicative and authentic language tests put forward by Morrow (1979). After communicative competence was proposed by scholars such as Widdowson (1978; 1979; 1983), Savignon (1972; 1983), and Canale and Swain (1981), the existing testing system started to be criticized severely. These scholars perceived language use as a dynamic process taking place in a situational context. In other words, they advocated asking learners to engage in the language use rather than the language itself by designing a context and evaluating their real performances. This communicative view pushed the language testers to pay regard to the discourse and sociolinguistic issues while designing the tests (Bachman, 2000).

Although the language testing field has come a long way in terms of employing communicative tests, implementation of these tests is still a hotly debated issue, especially to test learners' speaking skills.

Testing Speaking

Testing speaking has been a controversial issue in the field of language testing since there are a number of factors, that should be taken into consideration by testers such as "construct definition, predictability of task response, interlocutor effect, the effect of characteristics of the test-taker on performance, rating-scale validity and reliability, and rater reliability" (O'Sullivan, 2008, p.1). Moreover, not only test developers play a role in assessing learners' speaking skill, but also the other participants such as test-takers, interlocutors and assessors are highly important to the test discourse (Alderson & Bachman, 2004). Therefore, it is an accepted view

that assessing speaking skill is more difficult compared to testing the receptive skills of listening and reading in particular.

The views that testing should have a positive washback on teaching and learning and that testing and teaching are interwoven both influenced the dimension of testing speaking skill. Since there is an emphasis on performance-based speaking activities in language classrooms, it is also possible to see the implementation of performance tasks in speaking exams. For this reason, while in the past, learners' speaking skills were assessed through discrete-point tests, currently performance-based tests, which allow more interaction and negotiation of meaning are implemented.

It is suggested that performance tasks that test-takers engage in should represent the real-life situations that they may confront with and these tasks should enable test-takers to have a natural conversation (Weir, 1990). This view entailed the use of various types of oral interviews for the purpose of testing learners' speaking skills.

Types of Oral Interviews

Oral interviews can be categorized into three in terms of the number of test-takers that they involve; individual, paired and group tests. An individual test, which is also called as candidate-examiner test, allows only one test-taker and an examiner to have a conversation, whereas paired and group tests allow at least two test takers to interact with each other in the presence of an interlocutor and an assessor. Hence, the conversational nature of speaking skill has recently launched the use of paired and group test trend (Van Moere, 2013).

Individual Interviews (Candidate-Examiner). Individual interviews have been widely used since the early 20th century for the purpose of assessing the speaking skills of language learners. These tests have been employed by many government institutions such as the Foreign Service Institute (FSI) and Defense Language Institute (DLI), and nongovernment institutions such as the Educational Testing Service (ETS) and the American Council on the Teaching of Foreign Languages (ACTFL) (Johnson, 2000). In a typical Oral Proficiency Interview (OPI) test, there is usually an examiner and a test-taker whose performance is assessed by the examiner as well as a second rater during the exam. While individual interviews had emphasized dictation and pronunciation before the World War II, interaction and performance gained importance later on (Fulcher, 2003).

According to the proponents of individual interviews (e.g., Cubillos, 2010; Kenyon & Malabonga, 2001), these tests provide an authentic test discourse and context since test-takers have the chance to get involved in a real-life conversation in these tests. Advocates of individual interviews further claimed that individual interviews have high validity in testing learners' speaking ability because these tests are widely used by notable institutions. However, the construct validity of individual interviews was brought to the table in critical discussions by several researchers. For instance, Fulcher (2003) stated that the fact that individual interviews are employed in a widespread manner does not mean that they have high validity as the widely use of them does not prove that individual interviews are the best way of assessing learners' speaking skills. For this reason, the interactional organization of individual interviews and natural conversation has been compared in many studies (e.g., Johnson, 2001; Lazaraton, 2002; Van Lier, 1989; Young & Milanovic, 1992; Young, 1995; Young & He, 1998). All of these studies have agreed that the discourse of

individual interviews does not resemble the natural conversation in terms of speakers' turn-taking (Okada, 2010). While in an ordinary conversation both speakers have equal rights to take the stage, there is an asymmetric contingency between the examiner and the test-taker in an individual test discourse (Green 2014; Van Lier 1989). In other words, examiners gain an edge over the test-takers and they determine who will take turns or on what topic they will discuss during the test discourse (Okada, 2010). Therefore, their findings suggested that OPIs may have weak construct validity.

Paired speaking tests. The move towards a more communicative approach to language teaching has increased the use of pair work activities since the 1970s (Taylor, 2011). Considering that testing is an indivisible part of teaching, this approach has influenced not only the language teaching, but also the language testing field to a great extent (Ducasse & Brown, 2009) and the use of paired and group speaking tests both in classroom and assessment contexts has been more popular (Galaczi, 2014).

Paired speaking tests were first introduced by Cambridge ESOL in the First Certificate of English (FCE) examination in 1996 (Saville & Hargreaves, 1999; Taylor, 2000). In a typical paired speaking test, two test-takers interact with each other "in the presence of two examiners, one acting as an assessor and the other as an interlocutor" (Birjandi, 2011, p.171). To put it another way, test-takers engage in peer-peer interaction rather than candidate-examiner interaction in the paired format (Taylor & Wigglesworth, 2009). Much as the examiner takes part in the conversation to some extent, interaction is controlled by the test-takers, not by the examiner (Galaczi, 2008). Eygud and Glover (2001) asserted that peer interaction is more balanced in terms of quantity of talk, and topic initiation and extension when the

examiners play a minimal role in the test discourse, which turns out to be an advantage of paired speaking tests.

Several researchers suggested more potential benefits of paired format. To exemplify, Kormos (1999) claimed that paired speaking tests allow the test-takers to display their conversational management skills and manage the conversation better. Moreover, others have argued that test-takers approach these tests more positively (Egyud & Glover, 2001), since being tested with a partner decreases their communicative anxiety and stress (Ikeda, 1998; Norton, 2005; Saville & Hargreaves, 1999). Furthermore, it has been suggested that this format brings about positive washback in language classrooms as it leads to the growth of pair work and group discussion activities in language learning contexts (Ducasse & Brown, 2009; Saville & Hargreaves, 1999; Van Moere, 2013).

On the other hand, some scholars in the field asserted that paired speaking tests are problematic as the interaction is jointly constructed (McNamara, 1997; Swain, 2001; Young & He, 1998). McNamara (1997) and Swain (2001) proposed their concern regarding the assessment issue suggesting that although test-takers co-construct the interaction during the test discourse, their performances are scored and interpreted individually. What is more, the jointly created nature of discourse brings forward a possible peer-interlocutor effect both on scores (e.g., Davis, 2009; Galaczi, 2008; Iwashita, 2001) and interaction (e.g., Lazaraton & Davis, 2008) in paired speaking tests. Therefore, factors such as proficiency, familiarity, personality, identity and gender, which are brought to the test discourse by each test-taker, may influence the interaction or the spoken output while implementing these tests (Van Moere, 2013).

Group tests. In a typical group test, three or four test-takers interact with each other in the presence of an examiner; however, the examiner does not take part in the conversation (Sandlund et al., 2016). In a group test, the main aim is to elicit discussion. For this reason, topic cards and pictures are used to create a discussion and to enable test-takers to interact with each other in these tests (Hasselgren, 2000; Sandlund & Sundqvist, 2013).

Studies which have focused on group tests mostly investigated topic negotiation and turn-taking during the test discourse (Sandlund et.al, 2016). In these studies (e.g., Gan, Davison & Hamp-Lyons, 2009), it has been observed that group tests allowed test-takers to display their interactional skills better since test-takers need to have a control of the interaction to maintain the conversation. Greer and Potter (2012) stated that being tested in a group poses a challenge in turn-taking management for test-takers. Moreover, interlocutor effect on the test discourse and test-takers' performance is another possible problem of these tests. Some studies (e.g., Gan 2010; Nakatsuhara 2011) have indicated that test-takers' proficiency levels in a group test influence the interactional patterns which occur during the test discourse.

In this section, testing speaking and types of speaking tests in the field of language testing were discussed and in the next section, the concepts of interactional competence, interactional resources and interactional patterns will be presented.

Interactional competence

Interactional Competence (IC) can simply be defined as the knowledge of communication rules within a specific context. In the field of SLA, there are a number of scholars (e.g., Hall, 1993, 1995; He & Young, 1998; Kramsch, 1986;

Young, 2008, 2011), who have discussed the concept of interactional competence. The concept was first put forward by Oksaar (1983) and in his interactional competence model; he focused on the factors which may affect the construct of interactional competence. He claimed that paralinguistic features, nonverbal behavior, and sociocultural norms, which he named "cultureme and behavioureme" as well as extraverbal behavior such as proxemics, time and space may all influence interactional competence (1983, pp. 247).

The impact of interactional competence on the language testing field began to be felt with the proficiency movement put forward by Kramsch (1986). She emphasized interactional competence by stating that teachers should improve their learners' proficiency to make them interactionally competent. Moreover, she proposed that

> interaction entails negotiating intended meanings, i.e., adjusting one's speech to the effect one intends to have on the listener. It entails anticipating the listener's response and possible misunderstandings, clarifying one's own and the other intentions and arriving at the closed possible watch between intended, perceived, and anticipated meanings. (p.367)

To put it another way, interaction cannot be achieved individually, but it requires the collaboration between the interactants (Kramsch, 1986).

Kramsch's (1986) definition of interaction was furthered with the notion of co-construction, which was proposed by Jacoby and Ochs (1995) later on. The co-constructed nature of interaction was addressed by He and Young (1998) as well and they suggested that "interactional competence is not an attribute of an individual participant, and thus we cannot say that an individual is interactionally competent;

rather we talk of interactional competence as something that is jointly constructed by all participants". (p.7) To put it more simply, interaction emerges with the help of both speakers; otherwise it cannot be regarded as interaction. Therefore, interactional competence requires negotiating meanings and creating intersubjectivity, which is the basis for interactional competence (Young, 2008). Intersubjectivity is a term which refers to "the conscious attribution of intentional acts to others and involves putting oneself in the shoes of an interlocutor" (Young, 2011, p. 430). In other words, it can be considered as an attempt to provide mutual understanding between interlocutors.

Another point discussed under interactional competence is the term discursive practices, which is also called interactive practices. Interactive practice was defined by Hall (1995) as "structured moments of face-to-face interaction—differently enacted and differently valued—whereby individuals come together to create, articulate, and manage their collective histories via the uses of sociohistorically defined and valued resources" (pp. 207–208). Young (2008) mentioned these practices as discursive practices defining those "recurring episodes of social interaction in context, episodes that are of social and cultural significance to a community of speakers". (p. 57) He and Young (1998) stressed the interactive practice, which was an indivisible part of interactional competence and they stated that interactional competence cannot be considered separate from the interactive practice; it is even constructed in it. Interactants need to interpret the interactive practice first and shape their conversation accordingly in order to provide mutual understanding during the talk.

While constructing discursive practices, the resources such as verbal, nonverbal, interactional and identity resources all have a role (Kasper & Ross, 2013).

Among these resources, interactional resources are distinctive as they differentiate interactional competence from communicative competence (Young, 2000). Interactional competence entails interlocutors to be aware of when they need to take turns, how they need to repair a trouble in a conversation or how they will manage a topic and employing interactional resources serve these purposes.

Interactional resources

Interactional competence was defined by Young (2008) as "a relationship between participants' employment of linguistic and interactional resources and the contexts in which they are employed" (pp.100). From his definition of interactional competence, it can be inferred that he emphasized the use of linguistic and interactional resources in particular contexts. The use of interactional resources is of great importance in terms of interactional competence, which will help the speakers maintain the flow during the talk.

There are several interactional resources that speakers can employ while engaging in a conversation. In his paper, Young (2000) suggested four interactional resources: understanding of sequences of speech acts that are associated with a particular discursive practice and the ability of participants to "construct a practice with a specific register"; apply a range of turn-taking strategies; and manage topic initiation, topic life, and topic choice (pp. 6–7).

Young (2008) furthered his suggestion of interactional resources later on and explained them in a more detailed way. He added three more resources and proposed that the seven resources which constitute the interactional competence are brought to the interaction by interactants. Table 1 presents the resources suggested by Young (2008).

Table 1

Resources (Adopted from Young, 2008, pp.71)

Identity resources o Participation framework: the identities of all participants in an interaction, present or not, official or unofficial, ratified or unratified, and their footing or identities in the interaction
Linguistic resources o Register: the features of pronunciation, vocabulary, and grammar that typify a practice o Modes of meaning: the ways in which participants construct interpersonal, experiential, and textual meanings in a practice
Interactional resources o Speech acts: the selection of acts in a practice and their sequential organization o Turn-taking: how participants select the next speaker and how participants know when to end one turn and when to begin the next o Repair: the ways in which participants respond to interactional trouble in a given practice o Boundaries: the opening and closing acts of a practice that serve to distinguish a given practice from adjacent talk

As is seen in Table 1, as well as interactional resources which comprise of speech acts, turn-taking, repair and adjacency pairs, interlocutors should have a good command of identity and linguistic resources in order to interact properly.

Furthermore, according to Markee (2008) there are three components of interactional resources:

• language as a formal system (including grammar, vocabulary, pronunciation);

• semiotic systems, including turn-taking, repair, sequence organisation;

• gaze and paralinguistic features.

The fact that Markee (2011) stressed accuracy and fluency of interlocutors as indicators of their interactional competence was criticized by Walsh (2011) and he stated that interactional competence should be considered in light of the relationship between the linguistic and interactional resources as Young (2008) suggested.

In recent years, the interactional resources employed by the test-takers in paired and group tests have been examined in some studies (e.g., Galaczi, 2008; Gan, 2010; Galaczi, 2013). In those studies, the researchers explored whether the interactional resources vary depending on the proficiency factor. For instance, Gan (2010) investigated the interactional resources employed by the test-takers in a higher-scoring and a lower-scoring group and compared them. The results of his study revealed that higher-scoring group got involved in each other's topics a lot, which was demonstrated with the test-takers' competition for the floor and with overlaps they caused during the test discourse. According to Gan (2010) test-takers' contribution to each other's topics indicates their endeavor to accomplish the task and their tendency to take control of the interaction that they engaged in. What is more, the members of the higher-scoring group supported the speaker with minimal responses and agreements when they were in the listener role. When it comes to the lower-scoring group in this study, they were not able to meet the requirements of the task, spoke with many pauses and the listener support was limited (i.e., lack of

minimal responses and agreements). Moreover, their topic extension was limited too since they did not engage in each other's ideas. Galaczi (2013) also explored the use of interactional features across different proficiency levels and found out that topic development, listener support moves and turn-taking management varied as the test-takers' proficiency level increased. More specifically, lower level test-takers' other-initiated topic extension moves and listener support was limited and overlap was observed less frequently. However, higher level test-takers engaged in their partners' topics more, supported the speaker when they were in the listener role and the frequency of overlaps increased with the proficiency level.

All in all, it is apparent that in a conversation interlocutors need to possess interactional competence, which requires them to employ interactional resources in order to create mutual understanding and achieve intersubjectivity (Hall & Pekarek, 2011).

Turn-taking. Turn-taking is one of the most important organizations of talk-in-interaction since talking one by one is needed not to cause overlapping and any communication breakdowns during the talk (Schlegloff, 2007). Interlocutors need to talk one at a time and they should be aware of when they will end their own turn and allow the next speaker to speak (Feldstein & Welkowitz, 1987).

Turn-taking management in talk-in-interaction involves three components: turn length, turn speed and turn domination (Ducasse & Brown, 2009; Ducasse, 2010; Watanabe & Swain, 2007). Turn length, which refers to the frequency of turns and utterances in a conversation, has been a topic of several studies (e.g., Csépes, 2009; Ducasse & Brown, 2009; Norton, 2005; Taylor, 1999). In these studies researchers determined the turn length by examining the number of turns and

utterance length. Additionally, turn speed refers to how fast interlocutors respond to each other. For instance, Ducasse and Brown (2009) identified turn speed by analyzing the speed of interlocutors' responses to each other in their study. Moreover, turn domination was described by some scholars as a feature which shows how interlocutors struggle to stay on the stage (e.g., Ducasse & Brown, 2009; Dimitrova-Galaczi, 2004).

Another interactional resource which is an integral part of interactional competence is repair, which will be discussed next.

Repair. As Faerch and Kasper (1983) put it, during a conversation interlocutors may encounter communication problems; in situations of which they need to modify what they intend to say using repair strategies to be able to send a comprehensible message to the other speaker. Repair addresses "recurrent problems in speaking, hearing, and understanding." (Schegloff, Jefferson & Sacks, 1977, pp.361). In addition, it helps interlocutors organize and maintain the conversation they engage in. In the literature, several types of repair have been identified such as self-initiated self-repair, self-initiated other repair, other-initiated self-repair, other-initiated other repair, repetition, paraphrase, confirmation checks, clarification requests and comprehension checks (Drew, 1997; Nagano, 1997; Schegloff et al., 1977; Schegloff, 2000).

Self-initiated self-repair can be observed in a situation that speaker 1 cannot find the right word, but after a small pause s/he finds it her/himself without a prompting from speaker 2 (Levinson, 1983). What is more, if speaker 1 makes an attempt to find the right word, but cannot find it and speaker 2 fills in it for speaker 1, it becomes self-initiated other repair.

Alternatively, repair can be initiated by the other speaker as well. For instance, in other initiated self-repair, speaker 1 may not hear what has been uttered by speaker 2 or there may be some misunderstandings between them. In such a situation, speaker 1 asks for a clarification using some prompts such as ‘What?’ or ‘Sorry?’ and speaker 2 repairs the trouble source repeating or paraphrasing what s/he has uttered. Furthermore, if the trouble source is repaired by speaker 1 in the same situation, it becomes other-initiated other repair.

Apart from turn-taking and repair, interlocutors should have the ability of managing the topic that they are talking about in order not to create any communication breakdowns during the talk, for which they need to have interactional competence.

Topic management. Topic management in talk-in-interaction requires interlocutors to open, extend, connect and close a topic (Wang, 2015). In order to maintain a conversation, interlocutors need to start a new topic which is called as topic initiation (Dimitrova-Galaczi, 2004), develop the topic which is referred to as topic extension (Dimitrova-Galaczi, 2004; Ducasse & Brown, 2009; Gan, 2010; May, 2011b) and know when they will close the topic. It is also essential to connect the topic for interlocutors. In other words, they need to contribute to each other’s topics during the talk. When both interlocutors engage in the topic to extend it, it means that they have the other-initiated topic moves. On the other hand, when only one interlocutor attempts to extend the topic, it indicates that there is a self-initiated topic move during the talk.

Topic management also includes “an awareness of how speakers deal with changes in a topic, how they maintain a topic, and how they repair the interaction

when a misunderstanding occurs" (Burns & Joyce, 2000, pp.94). In a conversation, interlocutors resort to topic shift, namely change the topic from one to another when there is a silence, when they want to focus on another topic or when they disagree with the other speaker (Maynard,1980). Topic shift is usually associated with turn dominance as the topic which is discussed during the conversation is decided upon by the dominant speaker in the end.

All of these interactional resources mentioned above: turn-taking, repair and topic management constitutes different types of interactional patterns such as collaborative, parallel and asymmetric depending on how they are employed by speakers. The extent to which interactional resources are used by the speakers or dominance of one speaker during the interaction determines the interactional patterns that take place during a co-constructed test discourse.

Task management. Task management is closely related to the task completion component of a rubric. While rating task completion skills of the test-takers, the raters take into account whether the test-takers have achieved the communicative goal of the task (Ellis, 2003).

According to Council of Europe (2001), learner competences play a vital role in accomplishing a task. While performing a communication task, learners should activate their competences and employ the general and communicative strategies in order to meet the requirements of the task and to monitor or repair the task successfully. Furthermore, interaction requires the interlocutors to involve productive and receptive skills. However, in order to manage the interaction process properly, as well as the productive and receptive skills, interlocutors should collaborate to create mutual understanding between them since this is a joint discourse.

It is proposed in the Common European Framework of References (2001) that when the learners carry out a role-play task which is referred to as series assessment, at lower levels what they can do should be emphasized so the focus should be the task achievement. On the other hand, at higher levels, the learners can be assigned the tasks which indicate some aspects of their proficiency in their performances. With this in mind, when the lower levels and higher levels are paired together to carry out a role play task, they may display different performances which may influence the management of the task. Since the high levels are more competent and can use the strategies more effectively, they can easily take control of the task, whereas the low levels have difficulty in accomplishing the task successfully as they cannot understand what the task requires them to do due to their low level of competence.

Interactional patterns

Interactional patterns, which are also referred to as co-construction patterns, indicate what kind of interaction interlocutors generate in a conversation. The main components of interactional patterns are mutuality and equality. While mutuality refers to how interlocutors create shared ideas and the extent to which they engage in each other's contributions (Damon & Phelps, 1989; Dimitrova-Galaczi, 2004), equality refers to how the work of interlocutors in a conversation is distributed and the extent to which they control the interaction between them (Damon & Phelps, 1989; Dimitrova-Galaczi, 2004; Van Lier, 1996).

Interactional patterns were first defined by Storch (2002) who observed four different types of interactional patters in ESL writing pairs. Drawing upon equality and mutuality, she named the patterns as collaborative, dominant/passive, expert/novice, and dominant/dominant. Storch's (2002) study was significant for the

follow-up studies as her definition of interactional patterns enabled the other (i.e., Galaczi, 2004; Galaczi, 2008; May, 2011b) researchers to find out the interactional patterns in speaking. Figure 1 shows how the interactional patterns are established in dimensions of mutuality and equality.

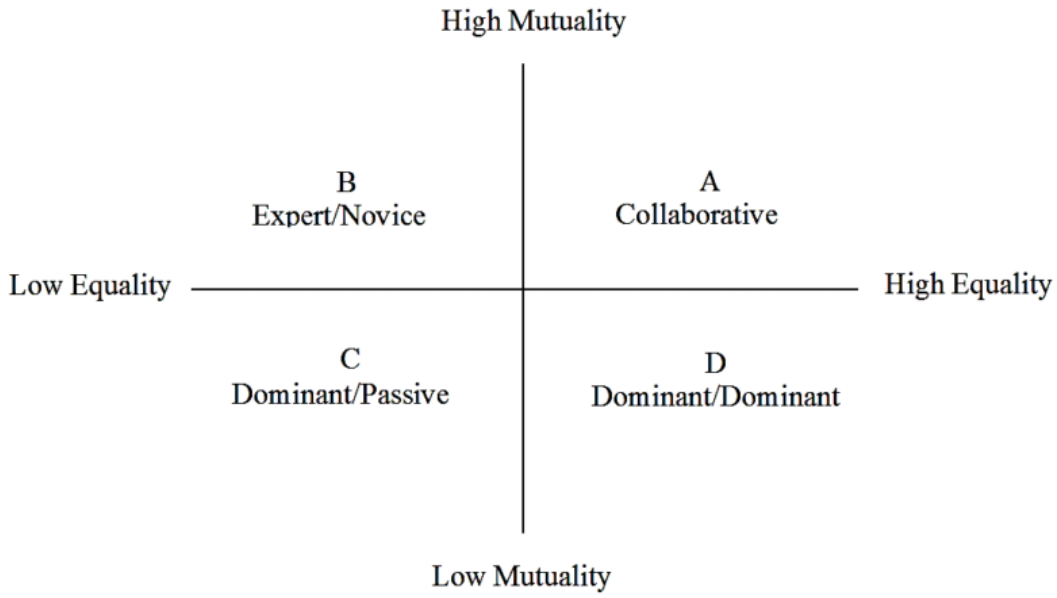


Figure 1. Interactional patterns in writing (adopted from Wang, 2015, p. 10)

Galaczi (2008) conducted a discourse-based study on the interaction between test-takers in a paired speaking test format. In her study, she examined 30 pairs and focused on the interactional features such as structural organization, turn- taking, sequencing, and topic organization. The aim of her study was twofold; she first examined the results of the analyses and determined four interactional patterns in talk-in-interaction by drawing on Storch's (2002) study. The patterns were termed as blend, collaborative, parallel, and asymmetric. Second, she investigated the relationship between interactional patterns and scores that raters assigned to test-takers. The results of her study yielded that while the test-takers who generated

collaborative pattern during the test discourse got the highest scores, the ones who created parallel pattern got the lowest scores.

According to Galaczi (2008) a collaborative pattern consists of high mutuality and high equality. To put it another way, it includes mutual feedback and shared ideas to a great extent and test-takers share the work in the conversation equally (Damon & Phelps, 1989). In Galaczi's (2008) study, test-takers in the collaborative pattern not only developed their own topics, but also extended their peer interlocutors' topics. On the other hand, in a parallel interactional pattern, interlocutors did not extend their partners' topics; rather they focused on and extended only their own topics, which mean that there was no contribution of partners to each other. Lack of collaboration in topic management indicates that this pattern involves high equality and low mutuality. When it comes to the asymmetric pattern, interlocutors were usually observed as one dominant and one passive in this type of pattern. This pattern presents low equality as there is an imbalance in the quantity of talk between interlocutors, whereas it allows moderate mutuality since interlocutors extended their partners' topics even if it is rare. Lastly, in a blend pattern, the researcher observed "discourse features of two interactional patterns" (pp.97). Figure 2 indicates the positions of interactional patterns in dimensions of mutuality and equality.

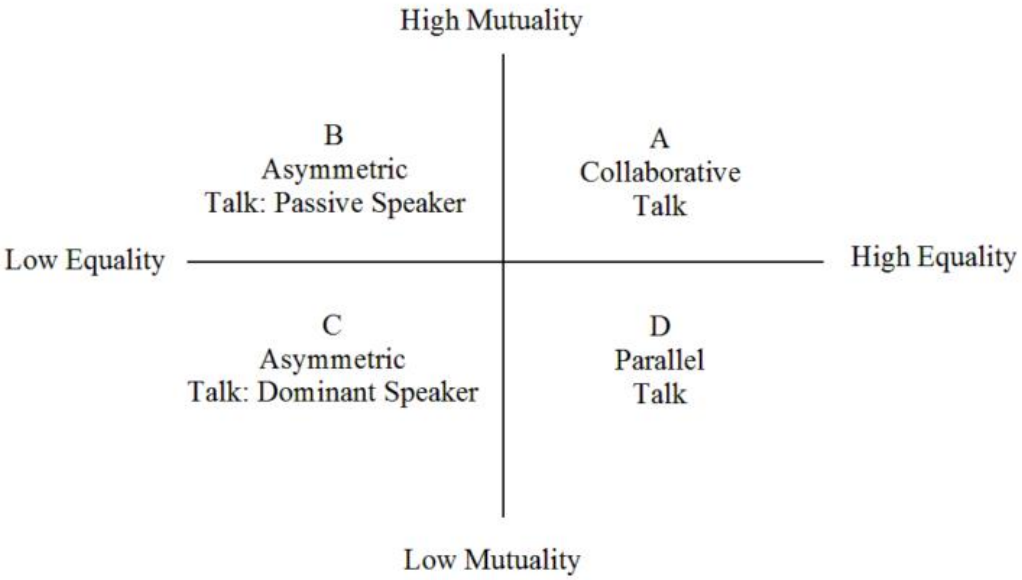


Figure 2. Interactional patterns in dimensions of mutuality and equality (Adopted from Wang, 2015, p. 11)

As seen in Figure 2, the interactional patterns that take place during a talk can vary depending on the mutuality and equality concepts. Although the interactional pattern model suggested by Storch (2002) was similar to the one put forward by Galaczi (2008) because they both use the concepts of equality and mutuality, Storch's model was developed for EFL writing pairs. However, Galaczi designed this model by examining the test takers in a paired speaking test format. For this reason, in the present study, Galaczi's (2008) model will be adopted.

To summarize, there are four types of interactional patterns: collaborative, parallel, asymmetric, and blend, which can occur during a talk and they are determined by drawing upon the interactional resources employed by interlocutors. Moreover, mutuality and equality are the concepts which should be taken into account while identifying the interactional patterns.

The next section will discuss the factors that influence the test discourse such as task effect and interlocutor effect which will be presented more elaborately.

Factors that affect test discourse

In this section of the literature review, the factors which affect test-takers' performances and/or their scores are presented taking the types of oral interviews as a whole. In other words, not only the factors which affect the discourse of paired speaking tests are mentioned, but individual and group tests are included as well.

The factors which influence the test discourse can be classified in two groups: task effect and interlocutor effect. While tasks themselves affect the test discourse, interlocutor effect involves proficiency, personality, familiarity and gender.

Task Effect

The effect of tasks used in oral interviews on test-takers' performances has drawn less attention by researchers in the field than the interlocutor effect. Furthermore, the impact of tasks has usually been investigated in the group speaking test discourse. As one of the researchers who examined the relationship between tasks and test-takers' performances, Van Moere (2007) compared various tasks such as picture tasks, consensus tasks and discussion tasks. The results of the study revealed that the picture task limited the quantity of talk produced during the test and length of turns. In contrast, the consensus task, which requires the test-takers to come to an agreement in a problem-solving or a decision-making situation, allowed the test-takers to interact with each other better compared to other two tasks and it was regarded as the best one among these three tasks. Another study was carried out by Nakatsuhara (2013) to explore the effect of a number of variables such as extroversion levels of test-takers, their proficiency levels and interlocutors on the

conversational patterns using different tasks in a group test. The findings of the study revealed that the proficiency level of the test-takers was a factor influencing the conversational patterns used in all tasks. However, it was claimed that extroversion levels affected the conversational patterns which occurred in more open tasks such as the ranking and free discussion tasks to a greater extent. Moreover, the effect of tasks in paired speaking tests has been investigated by Sandlund and Sundqvist (2013). Adopting a conversation analytic approach, they examined how test-takers and the examiner dealt with a discussion topic during the test. The results of the study suggested that test-takers and examiners showed different understandings of the task and these differences influenced the topic management in the test. In addition, Kley (2015) conducted a study and one aspect of her study focused on the task effect. She investigated whether the use of repair strategies, which are among interactional resources, differs based on the tasks used in paired speaking tests. The results of this study proposed that the types of repair strategies employed by the test-takers were usually similar across different tasks.

Interlocutor Effect

Test-taker characteristics such as proficiency, familiarity, gender and personality are considered as factors which affect test-takers' performances and their scores during the test discourse. In this section, studies conducted on these factors are covered separately.

Proficiency. Test-takers' proficiency level is considered by several researchers as one of the main factors which affects test-takers' scores and performances during the test discourse (e.g. Foot, 1999; Norton, 2005). For this reason, in the literature, there are several studies which investigated the proficiency

factor in paired speaking tests. For example, Iwashita (1996) explored if proficiency level of test-takers is a factor influencing the amount of talk and the scores assigned to test-takers in a paired speaking test. The findings of the study revealed that both low proficiency and high proficiency level test-takers performed better and produced more talk when they were paired with a high proficiency level test-taker. On the other hand, Csepes' (2009) study showed the opposite of what was found out by Iwashita (1996). He also examined the role of proficiency level on test-takers' performances, but he suggested that pairing test-takers with a high proficiency or low proficiency interlocutor does not influence the test-takers' performances and their spoken output. Moreover, Davis (2009) carried out a study on the effect of interlocutor proficiency on test-takers' performances, the amount of talk they produce during the exam and the scores assigned by the raters. In this quantitative study, test-takers were paired with interlocutors with both similar and different proficiency levels. The results of the Racsh analysis and paired t-test indicated that the proficiency level of test-takers slightly affect their performances, which is in line with Csepes' (2009) study.

While the aforementioned studies investigated the relationship between test-takers' proficiency levels and their performances / scores, there are fewer studies which examined the role of proficiency on interactional competence and interactional patterns that take place during the test discourse. Galaczi's (2013) study, for instance, explored how varying proficiency levels of test-takers correlate with the interactional resources such as topic management and turn-taking used in a paired speaking test. Furthermore, Gan (2010) investigated the relationship between test-takers' proficiency levels and the interactional features that they employed in a group test. The findings of both of these studies pointed out that the higher the proficiency level

of the test-takers was, the more they produced interactional resources during the test discourse.

Gender. Gender has been another factor which has caught researchers' attention in interlocutor effect studies. However, these studies have mostly been conducted in an individual test discourse in which a test-taker interacts with an examiner who is the interlocutor and the findings of these studies are variable. To exemplify, the results of Porter's (1991) study revealed that test-takers performed better when they were interviewed by a male interlocutor. In contrast, Porter and Shen (1991) conducted a study on the gender issue in oral exams with a similar research design with Porter's (1991) and this study put forward contrary results suggesting that the scores of the test-takers who were interviewed by a female were higher compared to its counterparts. Furthermore, O'Sullivan (2000) verified the Porter and Shen's (1991) finding with his study and proposed that test-takers performed better with a female interlocutor. On the other hand, O'Loughlin (2002) asserted that there is no gender effect on the scores of test-takers and on the test discourse. He further claimed that the gender effect found in the previous studies may be linked to the testing context which covers the tasks used and test-takers from different cultural backgrounds. Lastly, Brown and McNamara (2004) confirmed what O'Loughlin (2002) suggested and they stated that gender does not influence the test-takers' performances.

Personality. Personality is considered by several researchers in the literature as one of the factors influencing the test-takers' performances and scores. Berry (2004), for example, identified the test-takers as introvert and extrovert by giving them a questionnaire first before the group test that the test-takers took. After the test-takers were assigned scores by two raters during the exam, the researcher

interpreted the results. The findings of the study revealed that both introvert and exrovert test-takers scored better when the extrovert participants constituted the majority of the group. However, the scores of both introvert and extrovert participants were lower when the group consisted of mostly introvert test-takers. In the same vein to Berry (2004), Bonk and Van Moere (2004) investigated the role of shyness in group speaking tests. After the participants of the study discussed a topic in the test, the researchers conducted a survey to participants to look into their level of shyness. The results of this study indicated that outgoing participants gained an edge over shy test-takers during the group discussion, which was similar to the results of Berry's (2004) study.

On the other hand, Ockey's (2009) study proposed counter-views against both of these studies. He investigated whether test-takers' assertiveness and non-assertiveness affect their performances in a group test. The findings of his study showed that although assertiveness is regarded as a feature of extroversion, the groups consisting of mostly assertive participants did not affect the scores of either assertive or non-assertive test-takers. What is more, this result was verifed by Gan (2011). After giving the participants The Eysenck Personality Questionnaire, he determined test-takers' level of extroversion and explored the effect of extroversion on test-takers' oral performances. Similar to Ockey's (2009) results, the findings of Gan's (2011) study also suggested that there was no correlation between test-takers' level of extroversion and their oral performances.

Familiarity. It is evident in the literature that the effect of familiarity on test-takers' performances and their scores in paired speaking and group tests has been studied by many researchers. To exemplify, O'Sullivan (2002) conducted a study in which each of 32 Japanese test-takers got involved in a conversation with two

partners. While one of the partners was from the same college with the participants, the other partner was not familiar to them at all. The researcher interpreted the results of the statistical analyses in the way that familiarity factor affects the test-takers' both performances and their linguistic accuracy in paired speaking tests. When test-takers were paired with a partner with whom they are familiar, they worked better. Another study on the interlocutor familiarity conducted by Ying (2009) confirmed O'Sullivan's (2002) suggestion. She gave a questionnaire to participants to determine their acquaintanceship, right after they took the group oral test. After the participants replied the familiarity question 'Yes' or 'No, the researcher identified two groups with different familiarity levels; all-unfamiliar groups and mixed groups with two familiar members and an unfamiliar one. The results indicated that the test-takers were more comfortable when they were in the same group with familiar test-takers. In the study, Ying (2009) also examined the perceptions of participants of the familiarity factor. Surprisingly, the number of the participants who preferred to take the exam with strangers (30%) was much more than the ones who preferred to take the exam with familiar test-takers (20%). In addition, 50% of the participants offered no preference. Nevertheless, in the study carried out by Ockey, Koyama, and Setoguchi (2013), the participants of the study stated the opposite preferences. While 55% of the participants preferred to take the group oral exam with their classmates, 11% of them wanted to be tested with non-classmates. Moreover, in the study, the researchers examined if familiarity influences test-takers' scores during the test discourse and the results showed that there was no significant difference between the class-familiar and class-unfamiliar groups, which means familiarity effect on the participants' scores was not found out in the study. In the same vein to the study conducted by Ockey, Koyama, and Setoguchi (2013), Porter (1991a) also suggested

no acquaintanceship effect in his study in which 13 Arab learners were tested with familiar and unfamiliar examiners.

Four factors which are proficiency, gender, personality and familiarity were discussed in this section. As a summary, considering the studies conducted on the proficiency level factor on test-takers' performances and scores, it can be suggested that the results were indefinite. Some scholars proposed that proficiency is a factor which influences test-takers' performances and scores whereas some of them asserted the contrary. As to the gender factor, it has been investigated in candidate-examiner test discourse and there is no consensus on whether gender affects test-takers' performances in oral interviews in light of previous studies. For the personality factor, although some studies suggested that extrovert test-takers influence their partners' performances in a positive manner during a group test discourse, other studies claimed that extroversion does not have an impact on peer-interlocutors' performances at all. Lastly, whether familiarity is a factor which plays a role in test-takers' performances in oral interviews has been explored. While some scholars suggested that familiarity affects the test discourse, some of them asserted the opposite. Moreover, in some of these studies learners' perceptions of familiarity factor has been investigated as well. In light of these studies, the researchers obtained contradictory results; while some test-takers preferred to be tested with a familiar partner, some of them preferred a stranger.

Although proficiency and familiarity factors were investigated in the previous studies, no study has explored whether familiarity factor plays a role on interaction and few studies have explored the role of proficiency factor on interaction during a paired speaking test discourse. Therefore, in the present study this aspect will be addressed.

Conclusion

In this chapter, the literature related to the present study was presented in three sections. In the first section, after introducing language testing briefly, testing speaking and types of oral interviews were covered. Then the features of each oral interview format were explained and their advantages and disadvantages were covered separately with an extensive focus on paired speaking tests. In the second section, the literature on interactional competence and related concepts were provided. Lastly, after giving information about the factors which affect test discourse, summary of related studies on the factors such as task effect and interlocutor effect were presented. In the interlocutor effect part, factors such as proficiency, familiarity, gender and personality were covered separately with a particular focus on proficiency and familiarity factors. The next chapter introduces the methodology of the study.

CHAPTER III: METHODOLOGY

Introduction

The present study aims to investigate whether interactant related factors (i.e., familiarity and proficiency) play a role in a) EFL learners' use of interactional resources and b) the emergence of interactional patterns in paired speaking tests. To this end, in the study, interaction is operationalized as interactional patterns (i.e., collaborative, asymmetric, parallel and blend); and interactional resources (i.e., repair, turn-taking, topic management and task management). In this regard, the study addresses the following research questions:

1) How do the use of interactional resources and the emergence of interactional patterns vary in paired speaking tests with EFL test-takers

a) who have the same vs. different proficiency levels?

b) who are familiar vs. unfamiliar with each other?

This chapter includes five sections: the setting, participants, the research design, data collection and data analysis. In the first section the setting where the data were gathered is presented. In the second section, participants of the study are introduced. In this section the way the participants were selected from the archival data is described as well. In the third section, the research design is explained. In the fourth section, information about the data collection procedure is provided in detail. In the last section, the data analysis procedure is reported.

Setting

For the present study the data were retrieved from Bülent Ecevit University School of Foreign Languages which is located in Zonguldak, Turkey. The preparatory school at this state university provides one-year intensive English

courses to undergraduate students. At the beginning of each academic year, students take the proficiency exam and students who score under 60 are required to study at the preparatory school for one year. After students who will study at the preparatory program are identified, students take the placement test to be placed in appropriate classes in accordance with their scores. In the beginning, there are two proficiency levels at Department of Basic English of School of Foreign Languages; A1 and A2 according to the Common European Framework of Reference (CEFR). However, students are supposed to have the same exit level which is B1 at the end of each academic year. In order to prove that they are at B1 level, students take the final proficiency exam which measures their both receptive (reading and listening) and productive (writing and speaking) skills as well as their grammar and vocabulary knowledge. Students who score 60 or above pass the exam and become eligible to pursue their studies in their departments. The final exam constitutes 50% of the pass grade; the other half is determined by the average of midterms and quizzes which are conducted during the whole year and the portfolio which is kept by students throughout one academic year. The rationale for choosing this school is that it not only provides convenience sampling for this study, but also it is one of the few state universities which administers oral proficiency exam as a component of the proficiency exam and records these exams and keeps them in their archives. Furthermore, this school provided the archival data for the present study.

Selection of Sample

The sample of this study consisted of EFL students studying at Bulent Ecevit University. All students took the final exam along with an oral interview at the end of the 2015-2016 academic year, which was prepared at B1 level (CEFR). Since the exit level of the students was supposed to be B1 at the end of one academic year, all

students took the B1 level proficiency test. However, while some students scored high marks both in the proficiency exam and the oral exam, some of them scored low marks in both tests because of their low proficiency levels. Furthermore, at the institution where the data were retrieved, students were paired regardless of their proficiency levels for the oral proficiency exam, which leads to matching two high level students and two low level students together or pairing a high level student with a low level student. Thus, in this study, their scores were used to place the students into appropriate groups (i.e., low-low, high-high and low-high), which is crucial to answering the first aspect of the research question which investigates the role of students' proficiency on the interaction in paired exams. More specifically, as the first step, in order to identify and place the participants for low-high, high-high and low-low groups, their scores in the proficiency exam which measures their reading, writing and listening skills as well as their vocabulary and grammar knowledge, and their scores in the oral proficiency exam were examined by the researcher.

Oral proficiency exam constitutes 20% of the proficiency exam. Since both exams were examined separately during the selection of sample process, students' scores in the proficiency exam were evaluated over 80. Considering that students whose scores are 60 and above out of 100 pass the exam, the scores under 48 out of 80 were considered to indicate low proficiency level and the scores in the range of 48 and above were regarded as the indicator of high proficiency level. Moreover, the rubric used for grading students' oral performances comprises of 5 sections each of which is worth 4 points. Therefore, while the lowest score that students can gain is 5, the highest score is 20 in the oral proficiency exam. For this reason, while determining the high-high, low-high and low-low groups, scores in the range of 5 and 10 out of 20 were accepted as the indicator of low proficiency level and the

scores between the range of 15 and 20 out of 20 were regarded as the indicator of high proficiency level. After selecting the sample based on their scores, their oral exams were rated by the researcher once again and compared with their already assigned scores by two other assessors to achieve reliability. The reason why the videos were rated by a third rater was that the only criterion for the proficiency variable while grouping the students was students' scores. When there was discrepancy above two between their existing scores and the scores assigned by the researcher, those test-takers were not included in the study. In addition, one particular exam room was completely excluded and no participants were selected from this room because a high discrepancy appeared between the given scores by the assessors in this room and the scores assigned by the researcher. When test-takers scored under 48 in the proficiency exam, but scored in the range of 15 and 20, they were not accepted as low level. Similarly, when test-takers scored 48 and above, but scored between 5 and 10 were not regarded as high level. That is to say, while selecting the sample, the test-takers who met the requirements in both proficiency and oral proficiency exams were chosen as low levels and high levels in this study (See Appendix O and P for test-takers' scores). Consequently, 50 dyads that were paired as low-low, high-high and low-high were selected as participants of this study and this became the first cohort of participants to answer the first aspect of the research question.

As the second step, in order to answer the second aspect of the research question which investigates the familiarity variable, the researcher contacted all 100 students (50 pairs) in the first cohort and asked them if they were familiar with their peer interlocutors before the exam. In other words, they were asked whether their partners were their classmates or their friends. After getting their responses, nine

pairs out of 50 were detected as familiar and 41 of them were unfamiliar. While three of the familiar pairs were from the low-high group, six of them were from the high-high group. From the unfamiliar ones, another nine pairs were selected in order to compare them with the familiar ones. While selecting nine unfamiliar pairs, three unfamiliar pairs were selected from the low-high group and six of them were selected from the high-high group to create a homogenous group. Nonetheless, while analyzing the familiarity factor, their proficiency levels were not taken into consideration. To specify, the familiar pairs from the high-high group were not compared to the unfamiliar pairs from the same group, but they were compared to one another randomly. As a result, the researcher identified 36 participants (18 pairs) with familiar and unfamiliar partners for the second cohort (see Figure 3 for the distribution of the dyads).

First group (Proficiency Level)	Low-Low	High-High	Low-High
Number of the dyads	15	15	20
Second group (Familiarity)	Familiar		Unfamiliar
Number of the dyads	9		9

Figure 3: Distribution of the dyads

In brief, for the first cohort, in total, 100 participants were selected according to their proficiency levels. For the low-high group 40 participants were chosen, which was more than the low-low and high-high groups since this group was considered to be distinctive due to that one of the main aims of the study was to investigate whether pairing two different proficiency level test-takers play a role on

interaction in paired tests. For the second cohort, 36 of these 100 participants were selected after determining the familiarity factor by asking the students themselves.

The next section will present (a) students' scores in the oral proficiency exam and (b) students' scores in the proficiency exam which were used to select the participants.

Students' scores in the proficiency exam. At the end of each academic year, students at the institution where the data were retrieved take the proficiency exam to be able to pass the preparatory program which offers intensive English courses for one year. The proficiency exam consists of questions to test students' four skills which are reading, writing, listening and speaking as well as questions to gauge students' vocabulary and grammar knowledge. For the purpose of measuring students' reading and listening skills and their vocabulary and grammar knowledge, multiple-choice questions are asked. For the writing skill, students are asked to write a paragraph and for the speaking skill, an oral exam is administered. To be able to pass the program, students are required to score 60 or above in this exam.

Students' scores in the oral proficiency exam. As mentioned before, at the institution where the data were gathered, 20% of the proficiency exam consists of the oral proficiency exam. In the oral proficiency exam, students perform both an individual and a paired task and their performances are assessed by two raters. If there is more than 2 points discrepancy between two assessors, they negotiate themselves to reach a consensus. If they cannot make a consensus, they get counsel from the testing unit. When it comes to the rubric to evaluate students' performances, it comprises of five sections; fluency/pronunciation, vocabulary range, grammatical range/accuracy, spoken production and spoken interaction (See Appendix C). Due to that they are assessed through two different tasks, the rubric includes both a spoken

production section, which is for the individual task and a spoken interaction section which is for the paired task. Each section is 4 points which corresponds to 20 points in total.

Research Design

This descriptive study aims to examine the interactional resources such as turn-taking, repair, topic management and task management that the test-takers employed and the interactional patterns that occurred during a paired speaking test discourse. Test-takers' performances were recorded during the exam. In an attempt to observe test-takers' use of interactional resources and the emergence of interactional patterns during the test discourse, particular videos were transcribed by using the conventions suggested by Jefferson (2004) (See Appendix D for the transcription conventions) and analyzed by benefitting from the guidelines of Conversation Analysis (CA). For example, the components of CA which are turn-taking, repair, pauses and overlaps were analyzed in the scope of the present study. However, this study was not analyzed with a pure CA methodology; rather it was used as a framework during the analysis of videos. The next section will present the only data source 'video recordings' which provided the necessary data for this study.

Data Sources

In the present study, the only data source used was the video-recordings of the oral proficiency exam.

Video-recordings of the Oral Proficiency Exam

The oral proficiency exam consists of two phases at the institution where the data were gathered. Both of these phases are based on a performance test. To specify, in the first phase, students are required to perform an individual task. In the second phase, they are required to carry out a paired task. In the paired task, they are

assigned a role and make a conversation with their peer-interlocutors (See Appendix A for the role-play tasks). In this study, solely the second phase of the oral proficiency exam, which corresponds to approximately five minutes for each video was transcribed and analyzed (See Appendix B for duration of videos), since the study aims to examine the interactional resources and interactional patterns during the test discourse and paired speaking tasks are claimed to allow more interaction.

At Bulent Ecevit University, the School of Foreign Languages in Turkey each oral proficiency exam is video-recorded in case the test-takers object to their score and additional rating is needed. Additionally, all of the recordings are saved in the school archive. Therefore, for this study data were gathered from the videos recorded during the oral proficiency exam administered in 2015-2016 academic year. In other words, after receiving the required permission from the institution, the archival data were used.

In this study 50 videos of participants who were selected taking into consideration proficiency and familiarity factors were chosen in order to analyze test-takers' use of interactional resources such turn-taking, repair, topic management and task management; and the emergence of interactional patterns such as collaborative, asymmetric, parallel and blend.

Data Collection Procedure

After the required permission was granted from the administrators at Bülent Ecevit University, the School of Foreign Language, the archival data which belongs to 2015-2016 academic year were obtained from the institution. Getting the archival data, the selection of participant process began. First students' scores in the proficiency and the oral proficiency exams were examined. Then the students paired as high-high, low-low and low-high were determined paying regard to the range

which had been decided with the help of the researcher's advisor before. After identifying 100 participants for the first cohort, the researcher got in contact with all students and asked them whether they were familiar with their partners before the exam. Thereupon, from the first cohort, 18 pairs, 9 unfamiliar and 9 familiar, were determined as the second cohort.

When the selection of participants was completed, the researcher started to transcribe the second part (paired task) of the oral proficiency exam recordings. In total, 50 videos were listened to and transcribed using the conventions suggested by Jefferson (2004). Forthwith, all transcriptions were analyzed in order to identify the interactional resources (i.e., turn-taking, repair, topic management and task management) employed by the test-takers during the test discourse. Following this process, the researcher drew upon the interactional resources in order to assign the interactional patterns (i.e., collaborative, asymmetric, parallel and blend) which took place during test-takers' interaction with each other.

Data Analysis Procedure

In this study, data analysis procedure included three stages: (a) identifying the interactional resources, (b) determining the interactional patterns, and c) obtaining the descriptive results on SPSS. Table 2 indicates the interactional resources and interactional patterns analyzed in this study.

Table 2

Interactional Resources and Interactional Patterns Analyzed in the Study

Interactional Resources	**Interactional Patterns**
1. Turn-taking	A. Collaborative
1.1. overlap	B. Asymmetric
1.2. turn length	C. Parallel
1.3. turn speed	D. Blend (Collaborative+Asymmetric)
1.4. turn dominance	E. Blend (Parallel+Asymmetric)
2. Topic management	
2.1. topic initiation	
2.2. topic extension	
2.3. topic shift	
2.4. topic closure	
3. Task management	
4. Repair	

After transcribing 50 video-recordings by using the transcription conventions suggested by Jefferson (2004), the first stage of the data analysis comprised of identifying the interactional resources such as repair, turn-taking, topic management and task management employed by the test-takers during the test discourse. During the analysis procedure, the subcomponents of the interactional resources were also included so as to gain better understanding of interactional resources. To exemplify, since turn-taking management in talk-in-interaction includes three components: turn length, turn speed and turn domination (Ducasse & Brown, 2009; Ducasse, 2010; Watanabe & Swain, 2007) and no-overlap interaction is required in order to have a

smooth conversation with the other speaker, turn-taking was considered along with overlaps, turn speed, turn length and turn dominance. In the study, in an attempt to measure turn length the test-takers' quantity of talk and the number of turns they took were taken into consideration, whereas in order to measure turn speed the pauses that the interlocutors had while responding to each other were analyzed as Ducasse and Brown (2009) did in their study. Moreover, turn dominance was determined by taking into account the test-takers' turn length and their listener support (i.e., using minimal responses and agreements) when they were in the listener role. When it comes to topic management, it was categorized as topic initiation, topic shift, and topic extension and topic closure as maintaining a conversation entails the interlocutors to know how to start, shift and develop a topic, and when they need to close it (Dimitrova-Galaczi, 2004). Furthermore, while analyzing task management in the videos, whether the test-takers helped each other perform the task correctly in order to meet the requirements of the task was taken into consideration by the researcher.

While examining the interactional resources, color-coding was performed. To elaborate, a specific color was assigned to each resource and the analysis was carried out accordingly. After printing the transcribed talk of three groups (i.e., low-low, high-high and low-high), their use of interactional resources was first examined. Since turn-taking was considered with its subcomponents; turn speed, turn length, turn dominance and overlap, they were all highlighted with the blue color when they were observed in the transcriptions. Similarly, as topic management includes topic initiation, topic shift, topic extension and topic closure, they were highlighted with the green color when they were observed during the analysis of the transcriptions. Additionally, task management and repair were highlighted with pink and yellow

respectively (See Appendix R for the example of color-coding). Following that, interactional patterns (i.e., collaborative, asymmetric, parallel and blend) were determined by drawing upon turn-taking, topic management and task management, and their subcomponents which have been identified in the first phase (see Table 3). For instance, when one test-taker took longer and/or more turns during the interaction, and s/he was better at topic and task management than her/his peer interlocutor, asymmetric pattern was assigned to this pair (See Appendix S, T, U, V and Y for the samples of interactional patterns). Although repair was analyzed with the other interactional resources in the first phase of the data analysis because it is important in terms of maintaining the conversation, it was not taken into account during the analysis of the interactional patterns because repair strategies do not correspond to any interactional patterns. However, turn-taking, topic management and task management indicate whether test-takers collaborate or work individually, or whether one of them is more dominant during the test discourse. While analyzing the interactional patterns, the researcher drew on Galaczi's (2004) description. Table 3 presents the analysis of the interactional patterns.

Table 3

Analysis of the Interactional Patterns

Colaborative	• Topic extension by two speakers • Listener support by two speakers • Satisfactory other-initiated topic moves • Balanced turn-taking • Satisfactory task management by two speakers
Parallel	• Limited topic extension by two speakers • Limited listener support by two speakers • Limited other-initiated topic moves, but mostly self-initiated topic moves • Balanced turn-taking • Limited task management by two speakers
Asymmetric	• Topic extension mainly by one speaker • Listener support mainly by one speaker • Limited other-initiated topic moves • Unbalanced turn-taking • Unbalanced task management

Table 3

Analysis of the Interactional Patterns (cont'd)

Blend (Collaborative+Asymmetric)	• Topic extension by two speakers • Listener support by two speakers • Satisfactory other-initiated topic moves • Unbalanced turn-taking • Satisfactory task management by two speakers
Blend (Parallel+Asymmetric)	• Limited topic extension by two speakers • Limited listener support by two speakers • Limited other-initiated topic moves, but mostly self-initiated topic moves • Unbalanced turn-taking • Limited task management by two speakers

After completing the second phase of the data analysis, descriptive results were acquired using SPSS. While naming the values on SPSS in low-low and high-high groups in which students were at the same proficiency level, whether the conversation was maintained with the help of one test-taker or both of them was taken into consideration in general. However, in the low-high group, whether the high level test-taker or the low level one contributed more to the talk was taken into account. Since the nine pairs were selected from the 50 videos determined for the

first cohort for the familiarity group, as indicated in Figure 3, their videos were already analyzed while exploring the proficiency factor. Nevertheless, while naming the values in familiar and unfamiliar groups, in general, their equality in the use of interactional resources was considered, different from the values in proficiency groups. For example, whether familiar and unfamiliar pairs were equal or unequal in terms of their turn speed was analyzed.

Last but not least, in order to achieve reliability in coding, the researcher followed two main steps. First, all transcriptions were recoded by the researcher once more two months later to be able to provide intra-coder reliability. Second, an independent researcher was invited to code an interactional pattern to randomly selected five videos in order to provide inter-coder reliability.

Conclusion

In this chapter, the methodology of the study was explained. After introducing the setting and the participants of the study, the research design was described. Following that, the data collection and the data analysis procedures were presented. In the next chapter, the analysis of the data and the results will be provided in detail.

CHAPTER IV: DATA ANALYSIS

Introduction

The present study aims to investigate whether interactant related factors (i.e., familiarity and proficiency) play a role in a) EFL learners' use of interactional resources and b) the emergence of interactional patterns in paired speaking tests. In this study, interaction was operationalized as interactional patterns such as collaborative, asymmetric and parallel and interactional resources such as repair, turn-taking and topic management. In this regard, the study addressed the following research questions:

1) How do the use of interactional resources and the emergence of interactional patterns vary in paired speaking tests with EFL test-takers

a) who have the same vs. different proficiency levels?

b) who are familiar vs. unfamiliar with each other?

In this chapter, the results of the data analysis will be presented in accordance with the research questions in two main sections. In the first section, the role of pairing two same and two different proficiency level test-takers in the use of interactional resources and the emergence of interactional patterns will be provided. The use of interactional resources and the interactional patterns will be covered in three groups: high-high, low-low and low-high and the groups will be compared to one another in order to gain insight into whether proficiency factor plays a role in the use of interactional resources and the emergence of interactional patterns. In the second section, the role of pairing two familiar and unfamiliar test-takers in the use of interactional resources and the emergence of interactional patterns will be

presented. In this section, similar to the first one, the use of interactional resources and the emergence of interactional patterns will be discussed in two groups: familiar and unfamiliar groups and they will be compared to one another.

Results

The Use of Interactional Resources and The Interactional Patterns in Paired Speaking Tests with EFL Test-takers Who Have the Same vs. Different Proficiency Levels

In order to explore whether the use of interactional resources and the emergence of interactional patterns vary when two same and two different proficiency level EFL test-takers are paired, first 50 pairs were selected based on their proficiency and oral exam scores. After transcribing the videos of the participants by drawing upon Jefferson's (2004) conventions (See Appendix D), participants' use of interactional resources which include turn-taking, topic management, repair and task management were examined as the first stage of the data analysis. Following that, the interactional patterns were determined by looking at the interactional resources employed by the test-takers in the paired speaking test. As the last step of the data analysis procedure, the frequencies of the interactional resources and interactional patterns were obtained using SPSS.

Interactional resources employed by high-high, low-low and low-high pairs. In this study, there are 50 pairs of test-takers which were grouped as high-high, low-low and low-high based on their scores in the proficiency and the oral exam. There are 15 pairs in low-low and high-high groups separately, which makes 30 pairs in total, whereas the low-high group consists of 20 pairs as this is the distinctive group. In an attempt to find out the interactional resources; turn-taking,

topic management, repair and task management, employed by the test-takers, the interaction between the test-takers in 50 conversations was analyzed.

Turn-taking in high-high, low-low and low-high pairs. In the data analysis procedure, turn-taking was operationalized as overlaps, turn speed, turn length and turn dominance, which were analyzed in three different groups: high-high, low-low and low-high.

Overlap. To start with, during the analysis of overlap whether the test-takers lapped over during their interaction with each other was taken into consideration. When overlap was analyzed, the findings revealed that in this paired speaking test, one of the test-takers, both of the test-takers or neither of them can cause overlaps in a conversation.

An example of overlap in high-high pairs which was indicated with the convention '[…]' is as follows:

Table 4

Excerpt 1: Sample of Overlap in a High-High Pair

Tuğba (high-level): Thank you. (2.0) Can you bring the bill?
[Can you bring the bill?]
Rafet (high-level): [OK. Just a second.] (1.0) And it is your.
Tuğba: (1.0) Oh it is very expensive! I can I ask I ask for
just meat and water. [It's expensive.]
Rafet: [Yes but] it's our price and you can see it in menu.

As can be seen in Excerpt 1, during the conversation held by Tuğba and Rafet, only one test-taker brought about overlaps. More specifically, Rafet was the one who caused overlaps continuously. He did not wait for his partner to finish her

talk and took his own turn. As a result, their talk lapped over during their interaction with each other.

Figure 4 below presents the distribution of overlaps in high-high, low-low and low-high pairs (See Appendix E.1, F.1 and G.1 for the descriptive statistics of overlap).

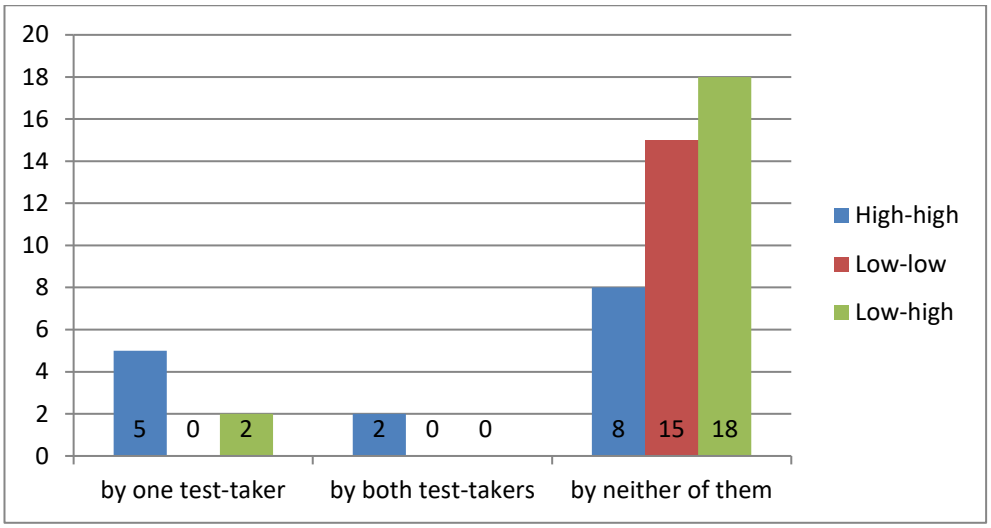


Figure 4. Distribution of overlaps in high-high, low-low and low-high pairs

Note: *In low-high pairs, overlap by high level test-taker=1; overlap by low level test-taker=1.

As it is clear in Figure 4, in eight conversations out of 15, there was no overlap between the two test-takers when they were paired as high-high. Nevertheless, while in two conversations, both test-takers caused overlaps during the interaction, in five of the conversations, only one test-taker brought about overlaps. Contrary to the findings regarding high-high pairs, in all 15 low-low pair conversations, none of the low level test-takers overlapped while they were talking to each other during the test discourse. What is more, in low-high pairs, in most of the conversations out of 20, there was no overlap between the test-takers and the two instances were done by either a low or high pair. More specifically, in two conversations, test-takers, in one of which the high level test-taker and in the other

one the low-level test-taker, overlapped whereas in 18 conversations, the test-takers had no overlaps between each other.

Overall, the findings revealed that overlap that takes place in paired speaking tests seems not to vary depending on pairing two different proficiency level test-takers because it is observed only when the high level test-takers are paired with other high levels. In low-low pairs, it does not take place during the test discourse. Additionally, although there is a high level test-taker involved, the low-high pairs do not have overlaps between each other since the low level test-takers do not attempt to support the speaker with minimal responses or agreements when they are in the listener role and they do not struggle for the floor.

Turn speed. Another feature of turn-taking is turn speed which refers to how fast the test-takers respond to each other. In the study, turn speed was analyzed in terms of the pauses that the test-takers had while responding to their peer interlocutors during their interaction with each other. The findings revealed that in the paired speaking test, peer interlocutors' turn speed can be the same or one of them can be faster than the other one.

A sample of turn speed in low-high pairs is as it follows:

Table 5

Excerpt 2: Sample of Turn Speed in a Low-High Pair

Demet (high-level): Hello welcome how can I help you what is your problem?
Bülent (low-level): (12.0) My problem is (7.0) don't feel
Demet: (2.0) OK (2.0) do you have any allergic (2.0) for any drugs?
Bülent: (4.0) No.
Demet: OK (2.0) I will give a (.) medicine for your (1.0)problem you (.) should use (1.0) two times in a in the day and you should go to a doctor for your problem because I don't know your problem
Bülent: (3.0) thank you very much.

As seen in Excerpt 2, Demet who was the high level test-taker was faster while responding to Bülent who was the low level one when their pauses are taken into consideration.

Figure 5 indicates the distribution of turn speed of test-takers in high-high, low-low and low-high pairs (See Appendix E.2, F.2 and G.2 for the descriptive statistics of turn speed).

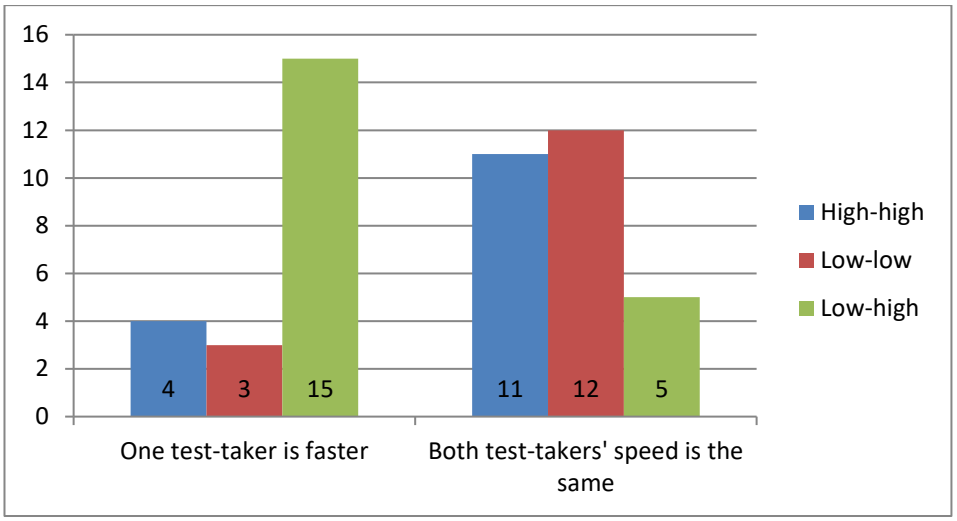


Figure 5. Distribution of turn speed in high-high, low-low and low-high pairs

Note: *In low-high pairs, high level test-taker is faster=14; low level test-taker is faster=1.

As shown in Figure 5, in high-high pairs, while only one of the test-takers was faster in four conversations out of 15, in 11 of them, both test-takers responded to one another at the same speed. Likewise, when a low level peer interlocutor was matched with another low level one, they responded to one another at the same speed in 12 conversations out of 15; still, in three conversations, one of the test-takers was faster than the other one. As illustrated in Figure 5 again, in low-high pairs, out of 20 conversations, in merely one conversation, the low level test-taker was faster than the high level test-taker. In contrast, in 14 conversations, the high level peer interlocutor was faster than the low level one. What is more, in five conversations held by low-high pairs, both test-takers talked to each other at the same speed.

All in all, the results yielded that turn speed seems to differ depending on the proficiency factor because when students are paired as high-high and low-low, in most of the conversations, both test-takers' speed is almost the same while responding to one another. However, in low-high pairs, the high level test-taker usually seems faster than the low level one.

Turn length. Turn length is another feature of turn-taking which was also examined in the study. It was observed that in a conversation, one test-taker can take longer turns or both test-takers can have the same turn length.

The following excerpt is a sample of turn length in low-high pairs:

Table 6

Excerpt 3: Sample of Turn Length in a Low-High Pair

Sinan (high-level): I can give (.) some advice you, (3.0) you (2.0) you book a meeting at a restaurant, you invite her you can invite her, and you buy flowers for her you you, (2.0)
Okan (low-level): She I don't flowers
Sinan: (2.0) If she don't like if she doesn't like flowers you can buy you should buy a clothes which she love, (.) For example
Okan: It'(h)s perfec(h)t.
Sinan: You should buy a jacket which (.) which she love maybe but if she doesn't like clothes you can write a poem, you can write lyric and you can sing a song for her.
Okan: OK thank you.

As can be seen in Excerpt 3, the high level test-taker Sinan took longer turns during his interaction with Okan who was the low level test-taker.

Figure 6 shows the distribution of turn length in high-high, low-low and low-high pairs (See Appendix E.3, F.3 and G.3 for the descriptive statistics of turn length).

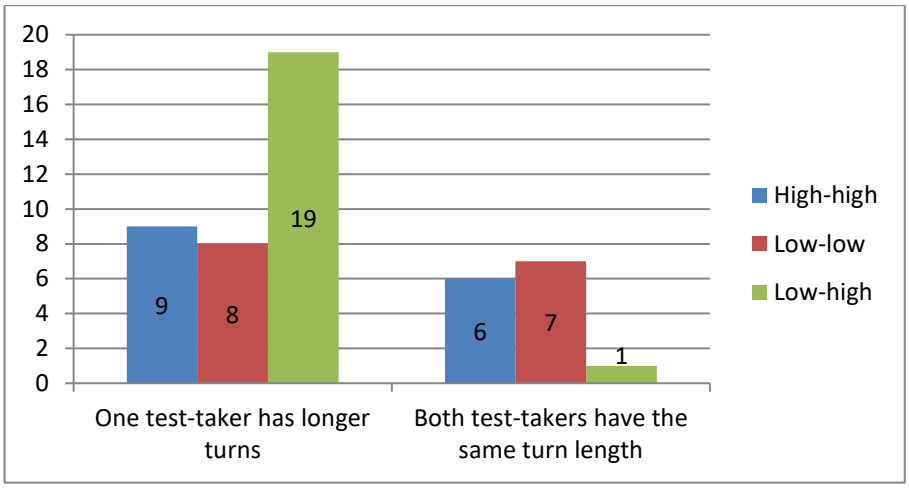


Figure 6. Distribution of turn length in high-high, low-low and low-high pairs

Note: *In low-high pairs, high level test-taker has longer turns=18; low level test-taker has longer turns=1.

As illustrated in Figure 6, in six of the conversations out of 15, both test-takers had the same turn length, whereas in nine of them, one of the test-takers took longer turns than the other one in the high-high group. Similarly, in low-low pairs, while in eight conversations, one test-taker had longer turns than the other one, in seven conversations, the length of turns taken by both test-takers was the same. As for low-high pairs, in most of the conversations (18 out of 20) high level test-taker had longer and/or more turns whereas in only one conversation, low level test-taker took longer turns. In addition, in only one conversation both test-takers' turn length was the same.

On the whole, the results of the data analysis revealed that turn length seems to differ depending on the pairing system. Although one of the test-takers usually seems to take longer turns in high-high and low-low pairs, they also have the same turn length in some conversations. However, when a low-high pair is matched, it does seem to play a role more since the high level test-takers seem to be always taking longer and/or more turns.

Turn dominance. Lastly, turn dominance was included in the data analysis as it is another subcomponent of turn-taking. In the present study, turn dominance has a close relationship with the turn length. The more or longer turns the test-takers take, the more dominant usually they are. Therefore, turn dominance of test-takers was basically determined by looking at how long and how many times the test-takers took turns and whether they responded to their peer interlocutors when they were in the listener role. In this sense, it emerged from the findings that in a conversation, one of the test-takers can be more dominant or there can be no dominance between the test-takers.

An example of turn dominance in low-high pairs is as it follows:

Table 7

Excerpt 4: Sample of Turn Dominance in a Low-High Pair

Hüseyin (low-level): What is your name?

Gülçin (high-level): (.) I am Gülçin, can I speak to the boss? (15.0)

Can I speak to the boss? (9.0)

((Because there is no attempt from Hüseyin, Gülçin tells the examiner that she wants to complete her task and speaks.))

Gülçin: I am so sorry my boss because I am I am ill therefore I can't go to the go to the work, (9.0) I am sorry for this but I can go to on Wednesday because I will go to the doctor thanks you so much for listening to me see you then my boss.

As it is obvious in Excerpt 4, Gülçin was the high level test-taker in this conversation and she took the turn to develop the topic although Hüseyin who was the low level test-taker did not attempt to extend the topic. Moreover, as Hüseyin did

not take turns to speak, Gülçin was more dominant in terms of turn-taking. She had longer and more turns than Hüseyin did.

Figure 7 indicates the distribution of turn dominance in high-high, low-low and low-high pairs (See Appendix E.4, F.4 and G.4 for the descriptive statistics of turn dominance).

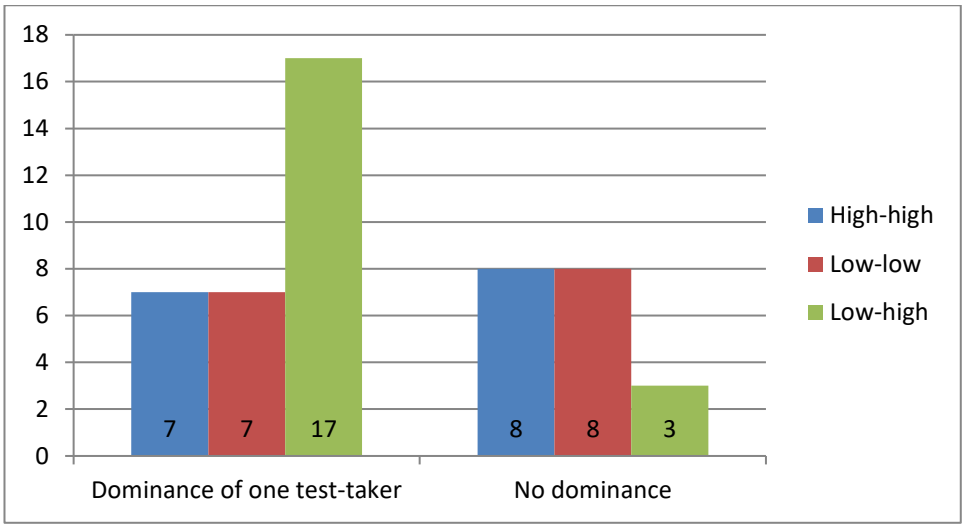


Figure 7. Distribution of turn dominance in high-high, low-low and low-high pairs

Note: *In low-high pairs, dominance of high level test-taker=16; dominance of low level test-taker=1.

As demonstrated in Figure 7, in high-high pairs, in seven of the 15 conversations, one of the test-takers was more dominant than the other one whereas in eight conversations, there was no dominance of one of the test-takers at all. Likewise, in the low-low group, although 'no dominance' was more frequent than the dominance of one of the test-takers, the number of the frequencies of two variables was highly close to one another. While in seven conversations out of 15, one test-taker was more dominant, there was no dominance in eight of the conversations. When it comes to low-high pairs, only in three conversations out of 20, there was no dominance of any test-takers and in merely one conversation the low level test-taker was more dominant. However, in the rest of the conversations which makes 16, high level test-takers were more dominant than the low level ones.

Overall, it can be claimed that, pairing the same level test-takers as high-high and low-low does not change the level of turn dominance in a paired speaking test since "dominance of one test-taker" was observed as frequently as "no dominance." On the other hand, when test-takers are paired as low-high, high level test-takers are definitely more dominant during their interaction with low level test-takers, which indicates that pairing two different level test-takers leads to more turn dominance in paired speaking tests.

Topic management in high-high, low-low and low-high pairs. In order to find out the extent to which test-takers could manage the topic they engaged in, topic management was operationalized as topic initiation, topic extension, and topic shift and topic closure.

Topic initiation. To begin, in the paired speaking test that the test-takers took, they were given role cards and who needs to start the conversation was written on their cards (see Appendix A). For this reason, during the analysis of topic initiation, the researcher examined if the topic was initiated by the test-taker that the task required, by the other test-taker that the task did not require or by the examiner.

The following excerpt is an example of topic initiation by the test-taker that the task did not require in low-low pairs:

Table 8

Excerpt 5: Sample of Topic Initiation by the Test-taker that the Task did not Require in a Low-Low Pair

Ilker (low-level): Hi.
(49.0)
Examiner: Please come on! If he doesn't start please you start the conversation.
Burcu (low-level): OK (5.0) My mother is (1.0) work every time, (3.0) I didn't have a babysitter but (2.0) I didn't have mother, so (3.0) I and my grandfather (1.0) with every time, so (5.0) I I had I had very friends but I am a shy child, (4.0) I didn't have a brother or sister, so (5.0) I am very (3.0) I am a angry and (1.0) cry. Every time.
Ilker: (2.0) What kind of games do you like?
Burcu: (4.0) I think I like babies.

As it is clear in Excerpt 5, since İlker did not start the conversation, the examiner asked Burcu to initiate the topic and then Burcu started to speak.

The following excerpt is an example of topic initiation by the examiner in low-low pairs :

Table 9

Excerpt 6: Sample of Topic Initiation by the Examiner in a Low-Low Pair

Aras (low-level): There there is there is a (1.0) problem.
Mehmet (low-level): OK. (.) Can you help you? (8.0) ((He realizes his mistake and looks at the examiner and his partner.))
Aras: Can you help me.
Aras: (8.0) I (2.0) I went to (4.0) a phar (2.0) phar (.) phar pharmacy.
Examiner: OK he is ill. Tell him your illness.
((The examiner reminds the test-takers their roles, and then they start the conversation once again.))
Mehmet: Yes. Hello
Aras: Hello
Mehmet: I want (2.0) aspirin because I am very cold.
Aras: (3.0) OK. (14.0)
Examiner: OK first of all ask him 'What is your problem?'
Aras: (4.0)What is your problem?

As can be seen in Excerpt 6, the conversation could not be started by the test-takers. They did not understand the task and their roles. For this reason, the examiner explained the the test-takers their roles and told them what they needed to say to initiate the topic.

Figure 8 illustrates the distribution of topic initiation in high-high, low-low and low-high pairs (See Appendix E.5, F.5 and G.5 for the descriptive statistics of topic initiation).

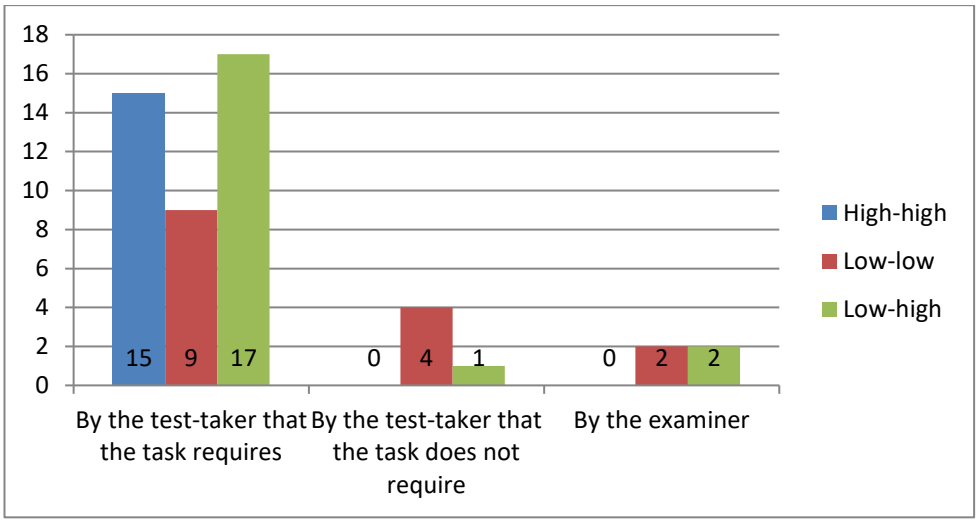


Figure 8. Distribution of topic initiation in high-high, low-low and low-high pairs

Note 1: *In low-high pairs, by the low level test-taker when the task requires=4; by the high level test-taker when the task requires=13. In addition, by the high level test-taker although the task requires the low level one=1; by the low level test-taker although the task requires the high level one=0.

As indicated in Figure 8, when test-takers were paired as high-high, the conversation was always started by the test-taker that the task required. Nevertheless, in the low-low group, in nine conversations out of 15, the topic was initiated by the test-taker that the task required. What is more, in four conversations it was started by the other test-taker that the task did not require and in two conversations, the topic was initiated by the examiner since neither of the test-takers could start the conversation. As can be seen in Figure 8 again, in 13 conversations out of 20, the task required the high level test-takers to start the conversation and they achieved it in the low-high group. In addition, in four conversations, the low level test-takers initiated the topic as the task required, yet in one of the conversations the low level test-taker could not start the conversation so the high level one initiated it. Moreover, in two conversations, the topic was started by the examiner since neither of the test-takers could start it.

All in all, it emerged from the findings that except a few instances especially in low-low pairs, pairing two same and two different proficiency level test-takers

does not have a role on their topic initiation. The test-takers seem to initiate the topic when the task requires them to.

Topic shift. Another subcomponent of topic management is topic shift. Although topic shift can be related to turn dominance, in this study, it was not considered as the indicator of turn dominance as the tasks did not always allow both test-takers to shift the topic because of the roles they were assigned to. Hence, topic shift was associated with only topic management in the present study. While analyzing the data, it was seen that the topic can be shifted by one test-taker, by both test-takers or by the examiner; however, it was also possible not to encounter a topic shift during the test discourse.

The excerpt below indicates a sample of topic shift in a conversation between two high level test-takers:

Table 10

Excerpt 7: Sample of Topic Shift in a High-High Pair

Peri (high-level): OK would you like to do another things?
Esra (high-level): For example, ıı ı (well) hide and seek maybe.
Peri: Hımm (Well). Kozlu beach?
Esra: Kozlu beach yes. Heh heh.
Peri: Heh heh.
Esra: (1.0) I think volleyball is good.
Peri: Okay. (2.0) I look, I look the weather and and it shows it might be rainy,
Esra: Yes.

Figure 9 shows the distribution of topic shift in high-high, low-low and low-high pairs (See Appendix E.6, F.6 and G.6 for the descriptive statistics of topic shift).

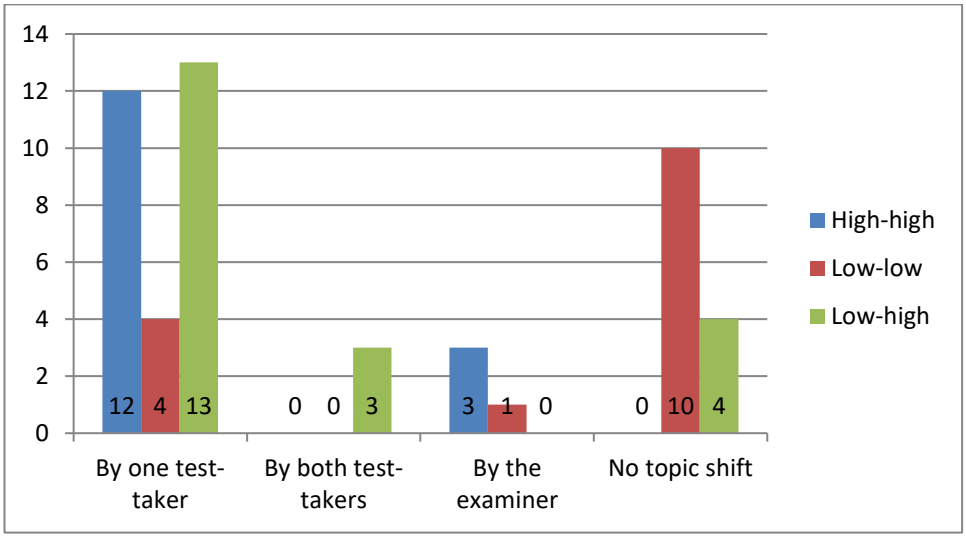


Figure 9. Distribution of topic shift in high-high, low-low and low-high pairs

Note: *In low-high pairs, by high level test-taker=8; by low level test-taker=5.

As demonstrated in Figure 9, in spite of that a high level test-taker was paired with another high level one, in 12 of the conversations out of 15, the topic was shifted only by one test-taker and in three of these conversations the topic was shifted by the examiner. Additionally, the results showed that in high-high pairs, there was topic shift in all conversations. Nevertheless, the topic was never shifted by both of the test-takers due to the nature of tasks assigned to them to perform. On the contrary, in most of the conversations made by low-low pairs, there was no topic shift. To specify, in 10 conversations out of 15, neither of the test-takers shifted the topic, whereas in four conversations, it was shifted by one of the test-takers. Moreover, in one conversation, the examiner interrupted the test-takers and the topic was shifted by the examiner. As to the low-high group, while in eight conversations out of 20, the topic was shifted by the high level test-takers, in five conversations, it was shifted by low levels. In addition, in three conversations, both the high level and

the low level test-takers shifted the topic and in four conversations there was no topic shift at all.

Overall, the findings indicated that proficiency seems to be a factor which plays a role on topic shift in paired speaking tests for the following reasons. First, when test-takers are paired as high-high, one of the test-takers seems to shift the topic resulting from the nature of the tasks assigned to the test-takers (i.e., the task requires one of the test-takers to ask some questions, whereas it requires the other test-taker to answer these questions, which hinders him/ her from shifting the topic.). On the other hand, even though the nature of the task remains the same, in low-low pairs, there is mostly no topic shift. Second, when low level test-takers are paired with high level test-takers, the topic usually seems to be shifted by the high level test-taker.

Topic extension. Another constituent of topic management is topic extension. In this study, topic extension was addressed from two perspectives: 1) who extended the topic, and 2) to what extent the topic was extended. During the data analysis, it was observed that the topic can be extended by one test-taker, by both test-takers and by neither of them.

An example of topic extension in high-high pairs is as follows:

Table 11

Excerpt 8: Sample of Topic Extension in a High-High Pair

Nesrin (high-level): You know, every every year, every spring term our school, our school planning a school picnic and I am responsible for some works about it and I will helped him. So I can I can ask to you some questions about the food. For example barbeque, what kind of food did you do you prefer? Meat or chicken?

Beyza (high-level): I think we can bring chicken or meat and vegetables or fruits. Maybe cake.

Nesrin: That's a good idea, eee (well)

(0.6)

Nesrin: But I am consider about the place. Maybe, it might be rainy the weather. So I can't I can't decide the school picnic place, where it is. Indoor or out? (.) What's your choice?

Beyza: (0.5) It's sunny, we will (0.5) we going to for instance maybe weather is rainy we will close place

Nesrin: OK yes.

(0.3)

Nesrin: Iıım (Well) school? (Unitelligible)

Beyza: (1.0) Maybe canteen?

Nesrin: Yeah, it's much more [(unintelligible)]

As can be seen in Excerpt 8, the topic was developed by both test-takers when they were paired as high-high.

Another example of topic extension in low-high pairs is as it follows:

Table 12

Excerpt 9: Sample of Topic Extension in a Low-High Pair

Bade (high-level): (9.0) Were you are a crazy boy?

Ayaz (low-level): (3.0) No (1.0) I I childhood is a (unintelligible).

Bade: (2.0) What kind what kind of games did you like it when you when were you child?

Ayaz: (1.0) My my favorite childhood played a (1.0) played a (2.0) football (1.0) player.

Bade: What kind of songs did you listen when you were a child?

Ayaz: (8.0) I don’t speak please ask questions

Bade: What kind of songs when you were a child?

Ayaz: (2.0) I don’t know.

Bade: (1.0) Now (1.0) do you meet your childhood friends?

In contrast, when a high level test-taker was paired with a low level one, the topic was usually extended by the high level test-taker and the low level one did not contribute to the talk much.

Figure 10 indicates the distribution of topic extension in high-high, low-low and low-high pairs (See Appendix E.7, F.7 and G.7 for the descriptive statistics of topic extension).

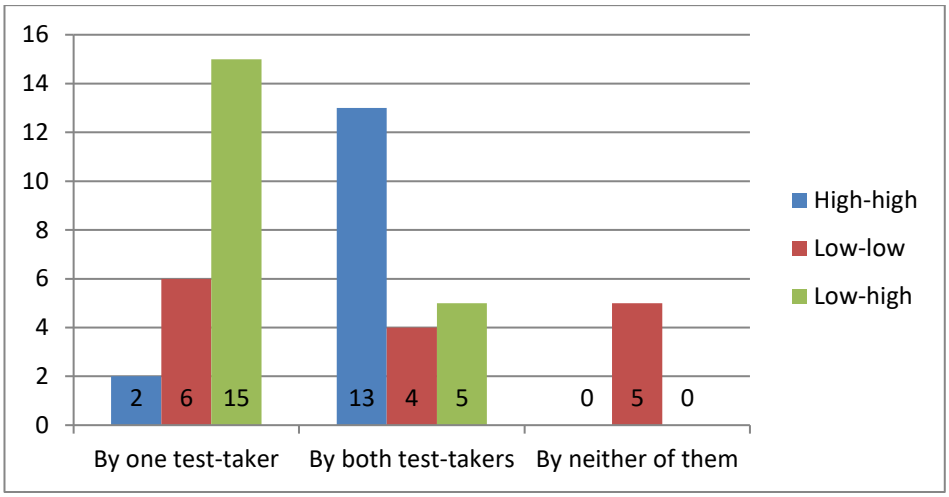


Figure 10. Distribution of topic extension in high-high, low-low and low-high pairs

Note: *In low-high pairs, by high level test-taker:15; by low level test-taker:0.

As shown in Figure 10, the topic is rarely extended by only one test-taker, which was observed in only two conversations out of 15; rather, it was extended by both test-takers in 13 conversations. However, in low-low pairs, the topic was extended by only one test-taker in six conversations out of 15. In comparison to high-high pairs, only in four conversations, both test-takers extended the topic and in five of them, neither of the test-takers extended it in the low-low group. Moreover, in low-high pairs, the topic was mostly extended by only high level test-takers and it was never extended by merely low level test-takers. To specify, in 15 conversations out of 20, only the high level test-takers extended the topic and in any conversations, the topic was developed by only low level ones. Nevertheless, in five conversations, low level test-takers extended the topic as well as the high level ones.

All in all, proficiency seems to be a factor which plays a role on topic extension because when test-takers are paired as high-high, the topic usually seems to be extended by both peer interlocutors and the topic is always developed during the test discourse. In contrast, when two low level students are paired together, there are instances of no topic extension and their collaboration seems limited while

extending the topic. Moreover, in low-high pairs, the topic always seems to be developed during the test discourse again, but it is mostly performed by the high level test-taker. For these reasons, high level test-takers seem better while developing the topic in paired speaking tests.

In an attempt to examine the second angle from which topic extension was addressed, the researcher identified three ranks: satisfactory topic extension, limited topic extension and no topic extension at all.

The following excerpt is an example of satisfactory topic extension in high-high pairs:

Table 13

Excerpt 10: Sample of Satisfactory Topic Extension in a High-High Pair

---

Yağız (high-level): I think we are planning together like this maybe we can do picnic on park you know,

Uras (high-level): Yeah that's a [great idea].

Yağız: [Green places] you know,

Uras: It's cool it's really cool.

Yağız: Yes. What about you? What do you think about it?

Uras: Yeah (1.0) I don't know I am I you know I am from Adana so we should make a mangal (grill) kebap (kebab),

Yağız: Yes of course maybe we can buy chicken or meat you know,

Uras: Oh I am gonna (going to) ask you something you know,

Yağız: Yeah of course.

As it is obvious in Excerpt 10, because both test-takers developed the topic collaboratively in high-high pairs, their topic extension was satisfactory.

Below is a sample of limited topic extension in the low-high group:

Table 14

Excerpt 11: Sample of Limited Topic Extension in a Low-High Pair

---

Ekin (low-level): (3.0) Hello

Umut (high-level): Hello

Ekin: (3.0) You are at school picnic. (11.0) What (2.0) time picnic school?

Umut: (3.0) Iıı (Well) I think (3.0) I start picnic sorry we are start picnic (3.0) twelve o'clock because we are hungry and (1.0) we are all together very activity for example volleyball and basketball and (2.0) tennis.

Ekin: OK. (9.0) Picnic is food,

Umut: (4.0) OK (3.0) I think I think we are eat we are eat Turkish meaning Turkish meaning because we are Turkish food is for example (1.0) mangal (grill) or (1.0) cızbız (grill). That's enough.

However, when a high level test-taker was paired with a low level test-taker, their topic extension was limited because the low level one did not work collaboratively with the high level one to extend the topic.

Figure 11 shows the distribution of topic extension in high-high, low-low and low-high pairs (See Appendix E.8, F.8 and G.8 for the descriptive statistics of topic extension-2).

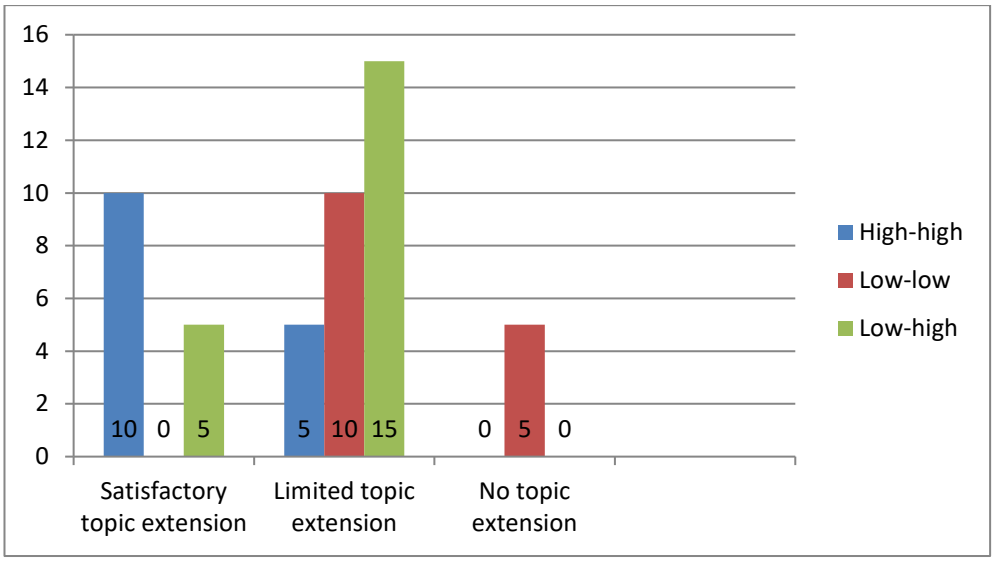


Figure 11. Distribution of topic extension in high-high, low-low and low-high pairs

As illustrated in Figure 11, the topics extended by high level test-takers were satisfactory in 10 of the conversations out of 15, whereas in five of them, the topic extension was limited. Still, in high-high pairs, although it was limited in some of the conversations, there was topic extension in all conversations. Thus, 'no topic extension' was never observed. Moreover, when Figure 11 is examined, it can be seen that in five conversations out of 15, the topic was never extended by the test-takers in low-low pairs. Additionally, in 10 conversations, the test-takers developed the topic, but it was limited. As can be seen in Figure 11 again, when a low level test-taker was paired with a high level test-taker, in only five conversations out of 20, topic extension was satisfactory and in 15 conversations, it was limited.

Overall, the findings regarding the second angle (i.e., to what extent the topic was extended) are in line with the first angle (i.e., who extended the topic) from which topic extension was approached. In this regard, it emerged from the findings that proficiency plays a role on the extent to which the topic is extended by the test-takers and while high level ones develop the topic in a satisfactory way, the topic extended by the low level ones is usually limited. Moreover, the low level test-takers

seem to influence the high level test-takers in a negative way while developing the topic. Seeing that low level test-takers cannot contribute to the talk much, topic extension in low-high pairs is limited again.

Topic closure. The last interactional resource which was analyzed under topic management was topic closure. The results of the data analysis yielded that in the conversations, the topic was ended by one of the test-takers, by both test-takers or by the examiner.

An example of topic closure by both test-takers in high-high pairs is as it follows:

Table 15

Excerpt 12: Sample of Topic Closure by Both Test-takers in a High-High Pair

Gizem (high-level): I think you can (.) go to doctor.
Turan (high-level): OK (3.0) how much is it?
Gizem: It's 5.99 dollar.
Turan: Heh heh OK OK please madam. ((He gives the money))Let's see you later. Thank you.
Gizem: Thank you.

A sample of topic closure by one test-taker in high-high pairs is as follows:

Table 16

Excerpt 13: Sample of Topic closure by One Test-taker in a High-High Pair

Kadir (high-level): OK. Which medicine can I use?
Seda (high-level): (3.0) Dicloron? [you should take] ((She smiles))
Kadir: [OK] thank you heh heh. How much is it?
Seda: Ten dollar.
Kadir: Thank you so much. Have a good day.
Seda: OK.

The following excerpt indicates a sample of topic closure by the examiner in low-low pairs:

Table 17

Excerpt 14: Sample of Topic Closure by the Examiner in a Low-Low Pair

Güneş (low-level): (4.0) What are symptoms?
(15.0)
((No answer from Akif))
Güneş: And you should see a doctor.
(19.0)
((They do not continue.))
Examiner: OK thank you the exam is over.

As it is clear in Excerpt 14, since the test-takers could not manage the conversation and did not attempt to maintain it, the conversation was ended by the examiner.

Figure 12 illustrates the distribution of topic closure in high-high, low-low and low-high pairs (See Appendix E.9, F.9 and G.9 for the descriptive statistics of topic closure).

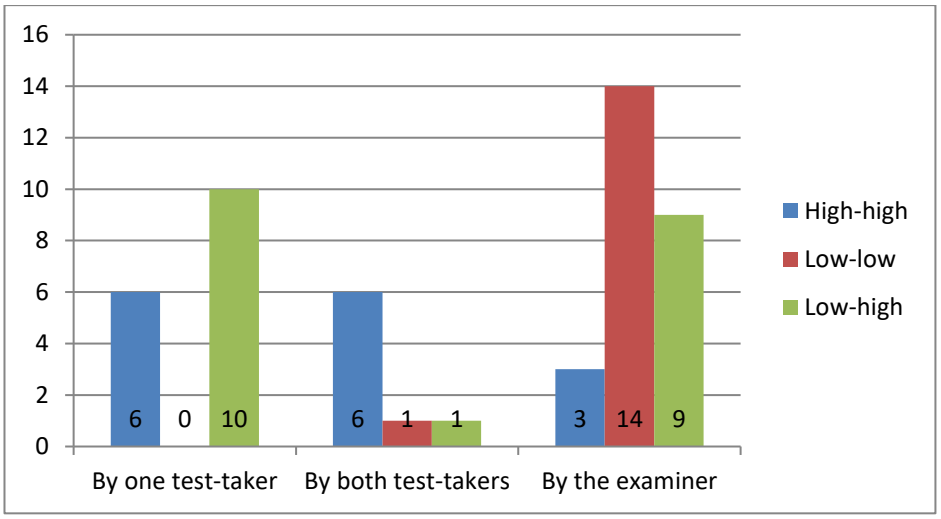


Figure 12. Distribution of topic closure in high-high, low-low and low-high pairs

Note: *In low-high pairs, by high level test-taker=7; by low level test-taker=3.

As shown in Figure 12, when a test-taker was paired with another test-taker who was at the same proficiency level with her/him, they usually ended their conversations themselves. More specifically, the topic was closed by one test-taker in six conversations out of 15 and by both test-takers in another six conversations. What is more, it was ended by the examiner only in three conversations. On the other hand, when test-takers were paired as low-low, only in one conversation out of 15, both speakers ended the topic themselves and in 14 conversations, the topic was closed by the examiner for the reason that the test-takers did not attempt to end the conversation. When it comes to low-high pairs, while in three conversations out of 20, the topic was closed by low level test-takers, in seven conversations, it was ended by high level test-takers and in only one conversation, both peer interlocutors closed the topic. Additionally, even though this group included a high level test-taker in each pair, in nine conversations the examiner ended the conversation owing to that the speakers could not close the topic themselves. However, in high-high pairs, the number of the conversations in which the examiner finished the conversation was three.

On the whole, proficiency definitely plays a role on topic closure in paired speaking tests. In high-high pairs, the topic is mostly closed by the test-takers, not by the examiner, whereas it is closed by the examiner in low-low pairs. Surprisingly, when high level ones were paired with low levels, the frequency of topic closure by the examiner and by the test-takers (mostly by high level test-taker) was very close to one another. Therefore, it can be claimed that low level test-takers cannot close the topic and they influence the high level ones in a negative way when they are paired together. However, when two high level test-takers are matched, they can close the topic and end their conversation.

Repair in high-high, low-low and low-high pairs.While analyzing repair in order to examine the effect of pairing two same and two different proficiency level test-takers together on it, it was found out that some of the test-takers repaired the trouble in a conversation when it was required whereas some of them did not repair even if it was required, which were taken into account when naming the values in SPSS. In addition, in some conversations no repair was performed since it was not required during the interaction, which was also taken into consideration.

Below is an example of repair applied when it was required in low-high pairs:

Table 18

Excerpt 15: Sample of Repair Applied When It Was Required in a Low-High pair

Şeyma (high-level): I sleep all night but (.) I don't feel well I want to a (.) drug but I have a allergic (1.0) allergic to aspirin (1.0) I can't eat aspirin drink as(h)pi(h)rin heh heh.
Bahar (low-level): OK. (12.0) You can drink aspirin OK, (6.0) ((In the meantime Şeyma recalls what the collocation of aspirin is))
Şeyma: Take aspirin! Drin(h)k aspiri(h)n tak(h)e aspiri(h)n heh heh
Bahar: Don't take aspirin
Şeyma: OK I know!

As can be seen in Excerpt 15, during the interaction the high level test-taker Şeyma recognized her mistake and was able to repair it. Her repair influenced the low level test-taker as well because Bahar also started to use the correct form afterwards.

The following excerpt illustrates repair which was not applied even though it was required in the low-low group:

Table 19

Excerpt 16: Sample of Repair Which Was Not Applied Even though It Was Required in a Low-Low Pair

Ilker (low-level): (3.0) Because very very
Büşra (low-level): (2.0) Handsome?
Ilker: ((He nods his head and smiles.)) Very handsome. (6.0) And who is your friends (were)?
Büşra: (1.0) Sorry?
Ibrahim: Friends, ((He murmurs and looks at Buşra's role card.)) (8.0)
Büşra: I don't understand.

As it is obvious in Excerpt 16, when test-takers were paired as low-low, one of them was not able to repair the trouble even though it was needed and it caused a communication breakdown.

An example of repair which was not applied even though it was required in the low-high group is as it follows:

Table 20

Excerpt 17: Sample of Repair Which Was Not Applied Even though It Was Required in a Low-High Pair

Kamile (low-level): Can you give me a menu?
Nihan (high-level): Yes of course. (2.0) Here you are.
Kamile: Eee (well), (4.0) I want to (1.0) cook (coke) and cheesecake please.
Nihan: Cheesecake.
Kamile: Yes.
Nihan: (1.0) Drinks?
Kamile: Cook (Coke)
Nihan: Cook?
Kamile: Yes
Nihan: OK anything else?
Kamile: No thanks.

Although it was not observed frequently, when students were paired as low-high, even though the low level test-taker did not repair the trouble, the high level one could maintain the conversation as can be seen in Excerpt 17.

Figure 13 demonstrates the distribution of repair in high-high and low-low pairs (See Appendix E.11, F.11 and G.11 for the descriptive statistics of repair).

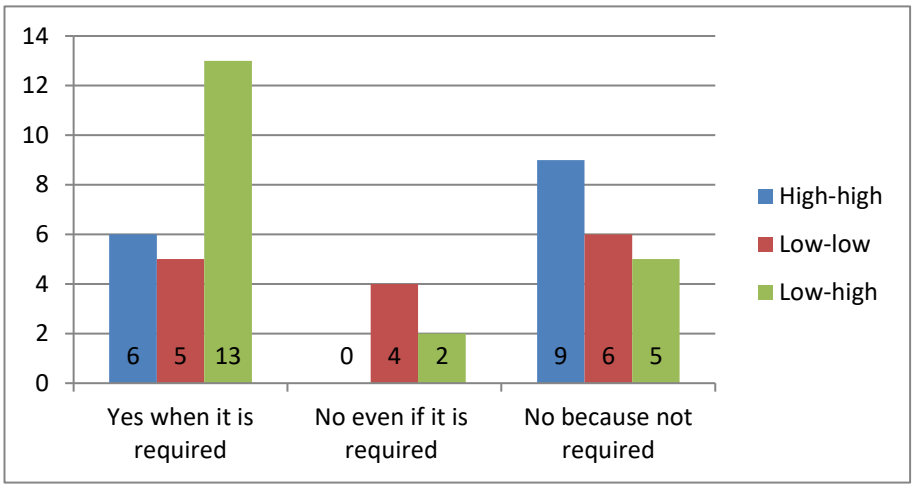


Figure 13. Distribution of repair in high-high, low-low and low-high pairs

Note: *In low-high pairs, yes by high level test-taker when it is required= 13; no by low level test-taker even if it is required=2.

As is indicated in Figure 13, in high-high pairs, in nine conversations out of 15, there was no repair because no communication breakdown took place and repair was not needed. Nevertheless, in six conversations, repair was required and the test-takers were able to repair the trouble during their interaction. Needless to say, there were not any test-takers who could not repair the trouble even though it was needed. When it comes to the low-low group, in five conversations out of 15, the peer interlocutors repaired the trouble when it was required and in six conversations, they did not resort to repair since it was not needed. Nonetheless, in four conversations, the test-takers did not repair the trouble in spite of that it was required during the interaction. While in some of these conversations, the fact that the speaker did not repair caused communication breakdowns, in some of them the test-takers were able to maintain the conversation. As it is clear in Figure 13 again, when high level and low level test-takers were paired together, high level test-takers were able to repair the trouble during their interaction in 13 conversations out of 20. On the other hand,

in two conversations, low level test-takers were not able repair the trouble even though it was needed. Lastly, in five conversations, both test-takers did not resort to repair since it was not required during the test discourse.

On the whole, the findings revealed that repair differs depending on the proficiency factor. When test-takers are paired as high-high, they mostly do not need to repair the trouble because there is usually no communication breakdown during the test discourse; however, when they are paired as low-low, they usually need repair during their interaction with each other since they do have troubles in their talk. Nevertheless, they sometimes cannot employ the repair strategies even though it is required. The result regarding repair in low-high pairs supports this finding because when low level ones are paired with high level ones, the repair is needed more and it is mostly applied by the high level test-takers.

Task management in high-high, low-low and low-high pairs. In this study, task management refers to whether the test-takers help each other perform the task correctly in order to meet the requirements of the task. During the analysis of task management in three groups: high-high, low-low, and low-high, it was found out that the task can be managed by one of the test-takers, by both test-takers or by the examiner. In addition, there can be no task management at all during the test discourse either because the test-takers are able to accomplish the task correctly without any help or because the test-takers are not proficient enough to manage the task.

A sample of task management by the high level test-taker in low-high pairs is as it follows:

Table 21

Excerpt 18: Sample of Task Management by the High Level Test-taker in a Low-High Pair

Bahar (low-level): I have a problem
Şeyma (high-level): (2.0) OK
Bahar: (9.0) I very ill
Şeyma: No that’s wrong I am ill!

As it is clear in Excerpt 18, the low level test-taker Bahar did not understand her role and started the conversation wrongly. Then the high level test-taker Şeyma managed the task and warned her about her role.

Another example of task management by the high level test-taker in low-high pairs is as it follows:

Table 22

Excerpt 19: Sample of Task Management by the High Level Test-taker in a Low-High Pair

Taner (high-level): Would you like to drink something or would you like to drink some meal?
(3.0)
Burak (low-level): Water(.) şey (well)
Taner: You are in the restaurant now ((He warns his partner.))
Burak: Yes restaurant Şey (4.0) Water cola

The following excerpt indicates a sample of task management by the examiner in low-low pairs:

Table 23

Excerpt 20: Sample of Task Management by the Examiner in a Low-Low Pair

Aras (low-level): Well OK. You can (2.0) you can (6.0) you can (1.0) go doctor (2.0) and (4.0) you should (3.0) aspirin buy a aspirin,
(13.0)
Examiner: OK give him a medicine.
Aras: Medicine (12.0)
Examiner: OK ask how much

As it can be seen in Excerpt 20, because the low level test-taker could not accomplish the task, the examiner helped him so that he could achieve the requirements of the task.

Figure 14 illustrates the distribution of task management in high-high and low-low pairs (See Appendix E.10, F.10 and G.10 for the descriptive statistics of task management).

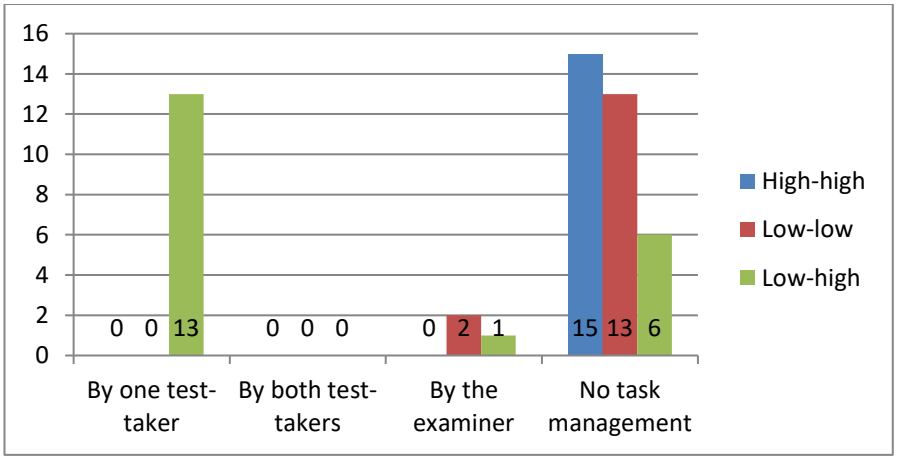


Figure 14. Distribution of task management in high-high, low-low and low-high pairs

Note: *In low-high pairs, by high level test-taker=13; by low level test-taker=0.

As indicated in Figure 14, in all 15 conversations, the test-takers did not manage the task seeing that there was no need for it. Since both test-takers in this group were high in terms of their proficiency level, they were able to accomplish the task correctly without any help from their peer interlocutors or the examiner. Hence, they did not have to manage the task to stay on it. As can be seen in Figure 14 again, in low-low pairs, in two conversations out of 15, the task was managed by the examiner as the test-takers could not meet the requirements of the task themselves because of their low proficiency level. In addition, low-low pairs did not manage the task in 13 conversations. However, the reason why they did not manage the task during their interaction was not the same with the reason in high-high pairs. In comparison to high-high pairs, the test-takers in this group were less proficient and they were not able to manage the task. In other words, they could not help each other stay on the task because of their proficiency level even though it was needed in some conversations. As for the low-high group, in six conversations out of 20, there was no task management at all during the interaction. On the other hand, while in only one conversation, the task was managed by the examiner, it was managed by the high level test-taker in 13 conversations, which makes a high percentage. However, the task was never managed by the low level test-taker.

Overall, the findings yielded that when test-takers are paired as low-low and high-high, task management (i.e., the need to fix the conversation to bring it on task) does not occur during the test discourse. The difference, however, between these two pairs are, while the high-high pairs do not need task management in paired speaking tests, the low level ones need it, but they cannot perform it. Nonetheless, when low level test-takers are matched with high levels, high level ones mostly manage the task during their interaction with each other.

Interactional Patterns generated by high-high, low-low and low-high pairs. In the present study, the second phase of the data analysis involved determining the interactional patterns generated during the test discourse by drawing upon the interactional resources which were presented in the previous section.

Interactional patterns in high-high pairs. To start with, Figure 15 demonstrates the distribution of interactional patterns in high-high pairs (See Appendix I for the descriptive statistics of interactional patterns in high-high pairs).

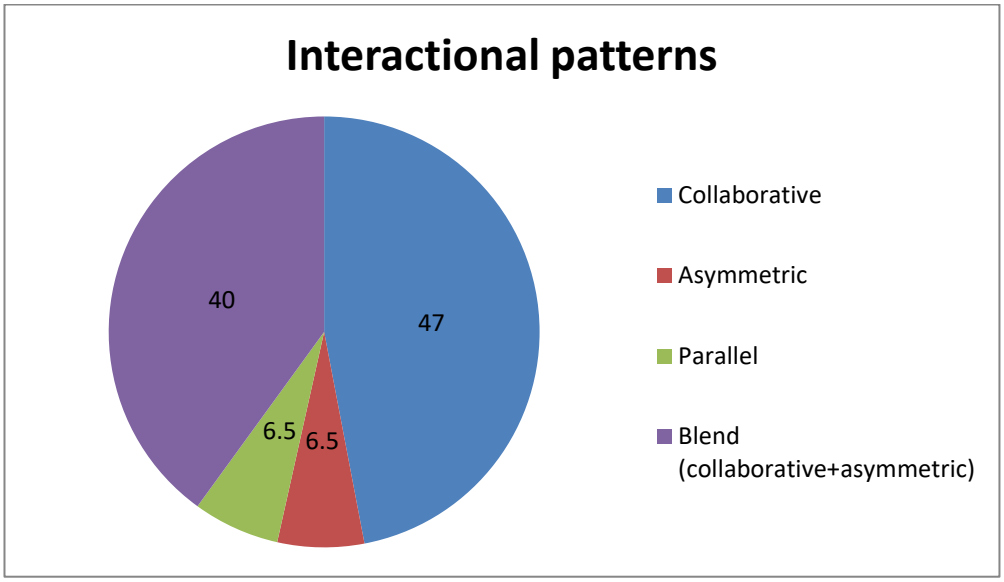


Figure 15. Distribution of interactional patterns in high-high pairs

As it is clear in Figure 15, in 47% of the conversations, the pattern generated by the test-takers was collaborative which includes high mutuality and high equality. As the test-takers in the high-high group were proficient enough, they were able to manage the topic collaboratively by contributing to each other's topics as well and the listener support (e.g., "Yes", "I agree with you") was good enough, which means that they used the minimal responses and agreements effectively . Moreover, there was balance between the test-takers in their quantity of talk. For this reason, most of the pairs generated a collaborative pattern during the test discourse. On the other hand, in 6.5% of the conversations, which corresponds to only one conversation out

of 15, the test-takers created asymmetric pattern which involves moderate mutuality and low equality. In this pattern, one of test-takers was more dominant in turn-taking, topic management and task management. The second pattern which was frequently observed in the high-high group was the blend of collaborative and asymmetric patterns with a percentage of 40. In this pattern, both peer interlocutors were balanced in terms of topic management; however, one of them was more dominant in turn-taking. Lastly, in the other 6.5% of the conversations, the pattern was parallel in which both test-takers worked individually while managing the topic, but their quantity of talk was almost equal.

Interactional patterns in low-low pairs. When it comes to the interactional patterns generated by low-low pairs, the results were a bit different from the ones in high-high pairs. Figure 16 indicates the distribution of interactional patterns in low-low pairs (See Appendix H for the descriptive statistics of interactional patterns in low-low pairs).

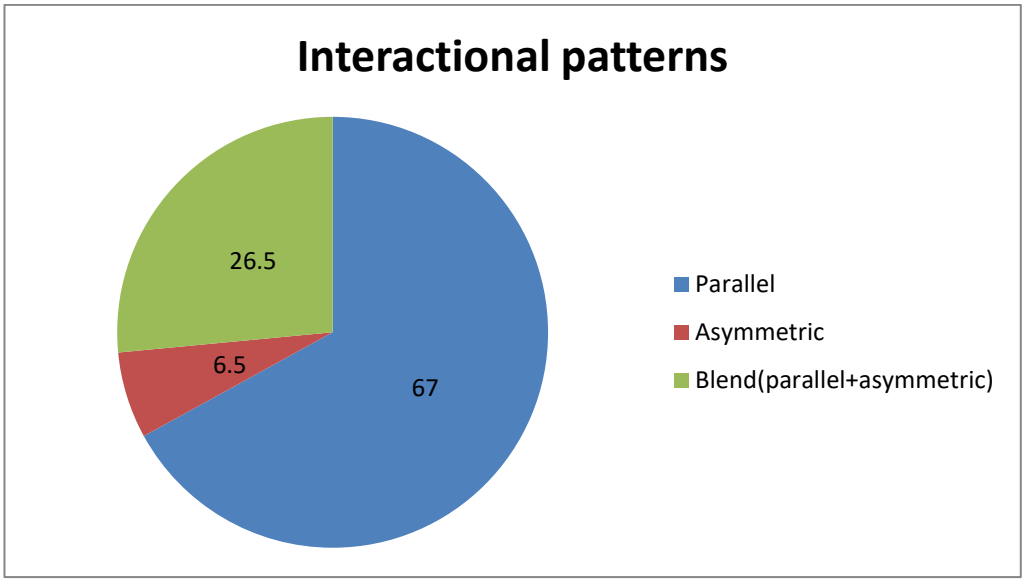


Figure 16. Distribution of interactional patterns in low-low pairs

As illustrated in Figure 16, in 67% of the conversations, the test-takers paired as low-low generated a parallel pattern owing to low mutuality and high equality

between the test-takers. Given that both test-takers were not proficient enough, they were not able to manage the topic and the task and the listener support (e.g., “Yes”, “I agree with you”) was highly limited. In other words, they were not able to contribute to the talk by developing the topic and show that they were listening to the speaker by using minimal responses and agreements. However, both test-takers were equal in their quantity of talk and topic expansion moves, which were again both limited. To put it differently, the test-takers in this group did not attempt to extend the topic much and they took short and few turns. Therefore, parallel pattern was assigned to those conversations. In addition, asymmetric pattern which involves moderate mutuality and low equality was observed at 6.5% which corresponds to only one conversation out of 15. The second most frequent pattern in the low-low group was the blend of parallel and asymmetric patterns with a percentage of 26.5. The reason why the blend of parallel and asymmetric patters together was assigned to the conversations in this group was that the interaction involved low equality and low mutuality. More specifically, the test-takers were not good at topic and task management and the listener support was limited and they were unequal in their quantity of talk and topic expansion moves as well.

Interactional patterns in low-high pairs. When it comes to the interactional patterns in the low-high group, four different interactional patterns were determined. Figure 17 indicates the interactional patterns in low-high pairs (See Appendix J for the descriptive statistics of interactional patterns in low-high pairs).

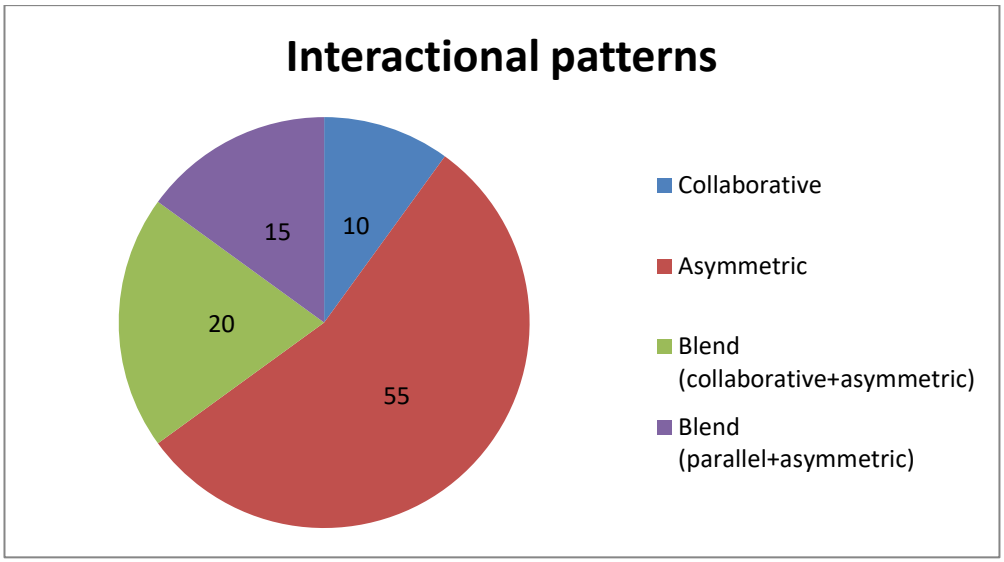


Figure 17. Distribution of interactional patterns in low-high pairs

As shown in Figure 17, asymmetric pattern which includes moderate mutuality and low equality was assigned to 55% of the conversations held by low-high pairs, which was the most frequent one in this group. Since one of the test-takers was high and the other one was low in terms of their proficiency, the high level test-taker was usually more dominant in turn-taking, topic management and task management. For this reason, asymmetric pattern was mostly assigned pattern in low-high pairs. Although it remains the least frequent pattern in this group, another pattern was collaborative detected in two conversations which corresponds to 10%. This pattern includes high mutuality and high equality. In other words, the test-takers were balanced not only in topic and task management, but also in turn-taking. Since the peer interlocutors in this group have different proficiency levels; one low level and one high level, the low level one was not able accommodate the high level peer interlocutor's pace. As a result, the high level one was more dominant in topic and task management and turn-taking in general. For this reason, collaborative pattern itself was assigned to only two conversations in this group. The second frequent pattern was the blend of collaborative and asymmetric patterns with a percentage of

20. Furthermore, another pattern generated by the test-takers paired as low-high was the blend of parallel and asymmetric patterns at 15%.

Table 24 below provides a summary of the findings regarding the interactional patterns in three groups: high-high, low-low and low-high.

Table 24

The Distribution of Interactional Patterns in High-high, Low-low and Low-high Pairs

	High-high pairs (%)	Low-low pairs (%)	Low-high pairs (%)
Collaborative	47	0	10
Asymmetric	6,5	6,5	55
Parallel	6,5	67	0
Blend (collaborative + asymmetric)	40	0	20
Blend (parallel + asymmetric)	0	26,5	15

As seen in Table 24, the interactional patterns differed depending on the pairing system (i.e., high-high, low-low and low-high) in paired speaking tests. The most frequent pattern generated in high-high pairs was the collaborative one whereas in low-low pairs, it was parallel and in low-high pairs, it was asymmetric. On the other hand, while the blend of collaborative and asymmetric patterns was the second most frequent one in high-high and low-high pairs, it was never observed by the test-takers in the low-low group, but they created the blend of parallel and asymmetric patterns.

The Use of Interactional Resources and The Interactional Patterns in Paired Speaking Tests with EFL Test-takers Who Are Familiar vs. Unfamiliar with Each Other

Among the 50 pairs examined in this study, only nine of them were familiar with each other and 41 of them were unfamiliar. However, only nine unfamiliar pairs were selected to be compared to the familiar ones. While selecting the unfamiliar pairs, their group was taken into consideration; however, the proficiency factor was not controlled.

Interactional resources employed by familiar and unfamiliar pairs. While analyzing if familiarity is a factor which plays a role in the use of interactional resources in paired speaking tests, mainly two categories were taken into consideration: if the familiar and unfamiliar pairs were equal or unequal (e.g., in turn length, whether the test-takers' quantity of talk was equal or not was explored). However, in some of the resources such as overlap, topic initiation and repair, some certain additions or changes were made (e.g., while examining overlap, whether there was overlap or not was explored, but the test-takers' equality was not taken into account).

Turn-taking in familiar and unfamiliar pairs. As mentioned earlier, while analyzing turn-taking, it was operationalized as overlap, turn speed, turn length and turn dominance in this study.

Overlap. To begin, while examining overlap in familiar and unfamiliar pairs, two categories were taken into account: 'there is overlap' and 'no overlap'.

For an example of overlap, please see Excerpt 1 (Table 4).

Figure 18 indicates the distribution of overlap in familiar and unfamiliar pairs (See Appendix K.1 and L.1 for the descriptive statistics of overlap in familiar and unfamiliar pairs).

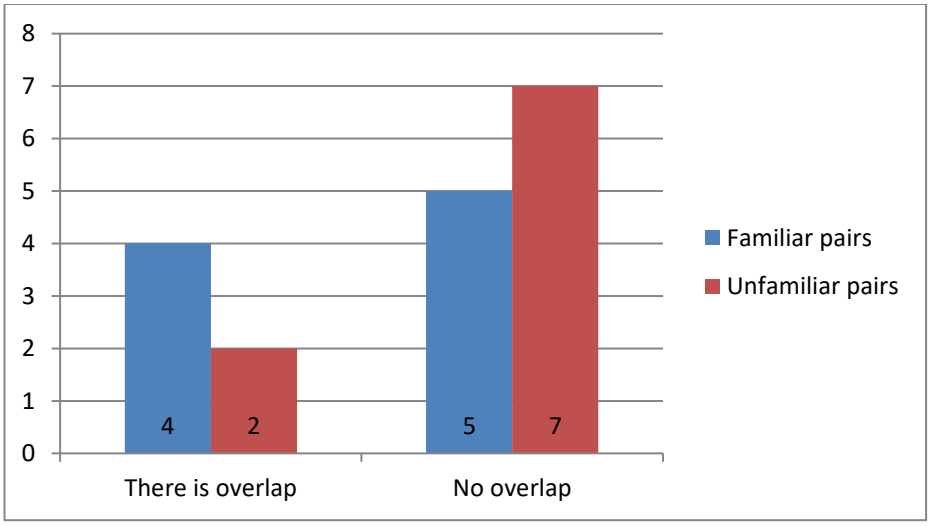


Figure 18. Distribution of overlap in familiar and unfamiliar pairs

As indicated in Figure 18, while the familiar pairs overlapped in four conversations out of nine, there was no overlap between the test-takers in five conversations. When it comes to unfamiliar pairs, in only two conversations out of nine, there was overlap between the test-takers; nonetheless, they did not overlap in seven conversations, which outnumber the frequency of overlaps in familiar pairs.

All in all, familiarity seems to influence the overlap that occurs in paired speaking tests. Even though lack of overlap was more frequent in both familiar and unfamiliar pairs, when two groups (i.e., familiar and unfamiliar) are compared, it is observed that familiar pairs cause more overlaps than unfamiliar pairs do.

Turn speed. While examining turn speed in familiar and unfamiliar pairs, whether test-takers' speed was equal or unequal while responding to each other was analyzed.

For a sample of turn speed, please see Excerpt 2 (Table 5).

Figure 19 illustrates the distribution of turn speed in familiar and unfamiliar pairs (See Appendix K.2 and L.2 for the descriptive statistics of turn speed in familiar and unfamiliar pairs).

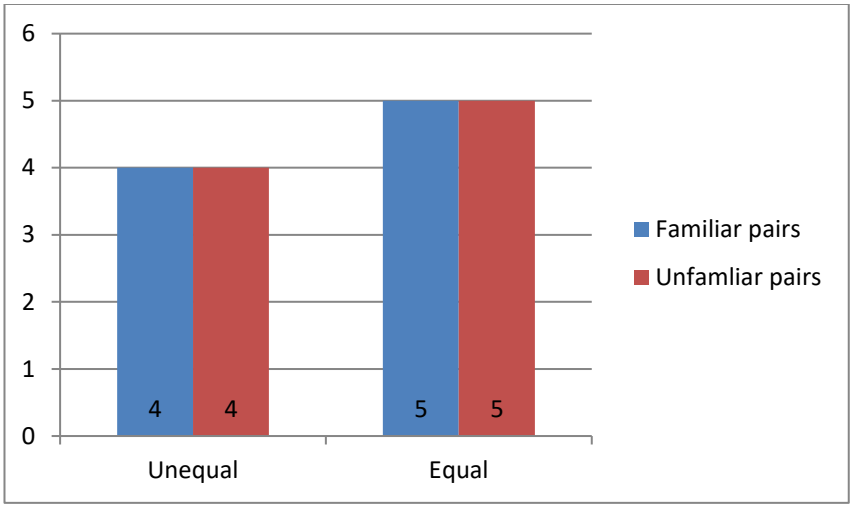


Figure 19. Distribution of turn speed in familiar and unfamiliar pairs

As shown in Figure 19, while in four conversations out of nine, one of the test-takers was faster than the other one, in five conversations they responded to each other at the same speed. In the same vein to the results of turn speed in familiar pairs, in four conversations out of nine, one test-taker was faster and in five of them, test-takers were equal in terms of their turn speed in unfamiliar pairs.

On the whole, the results indicated that familiarity does not have a role on the test-takers' turn speed since the frequencies of 'equal' and 'unequal' categories were exactly the same with one another in both familiar and unfamiliar pairs.

Turn length. Another subcomponent of turn-taking is turn length which was also examined in familiar and unfamiliar pairs. Again while analyzing turn length the researcher looked into if the test-takers' length of turns was equal or unequal.

For an example of turn length, please see Excerpt 3 (Table 6).

Figure 20 shows the distribution of turn length in familiar and unfamiliar pairs (See Appendix K.3 and L.3 for the descriptive statistics of turn length in familiar and unfamiliar pairs).

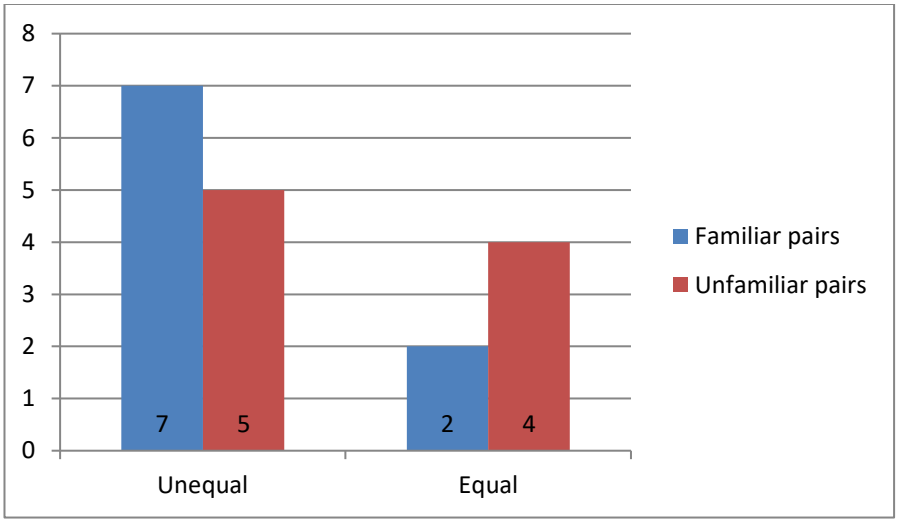


Figure 20. Distribution of turn length in familiar and unfamiliar pairs

As demonstrated in Figure 20, in familiar dyads, in seven conversations out of nine, one of the test-takers took longer or more turns, whereas in merely two conversations, their turn length was balanced. On the other hand, in unfamiliar pairs, in five conversations out of nine, one test-taker had more or longer turns and in four conversations, both peer interlocutors' turn length was equal. Therefore, in comparison to the familiar pairs, unfamiliar pairs were more balanced in terms of their turn length in this study.

Overall, the findings revealed that familiarity may be a factor which has a role on the test-takers' turn length because inequality in the turn length is more frequent in familiar pairs than in unfamiliar pairs. This finding indicates that in familiar pairs, one of the test-takers takes longer and more turns than in unfamiliar pairs.

Turn dominance. Lastly, while examining turn dominance in familiar and unfamiliar pairs, test-takers' equality in terms of their turn dominance was taken into consideration. Moreover, similar to the analysis of turn dominance in high-high, low-low and low-high pairs, the researcher took into consideration, in familiar and unfamiliar pairs, how long and how many times the test-takers took turns and whether they responded to their peer interlocutors when they were in the listener role.

For the example of turn dominance, please see Excerpt 4 (Table 7).

Figure 21 demonstrates the distribution of turn dominance in familiar and unfamiliar pairs (See Appendix K.4 and L.4 for the descriptive statistics of turn dominance in familiar and unfamiliar pairs).

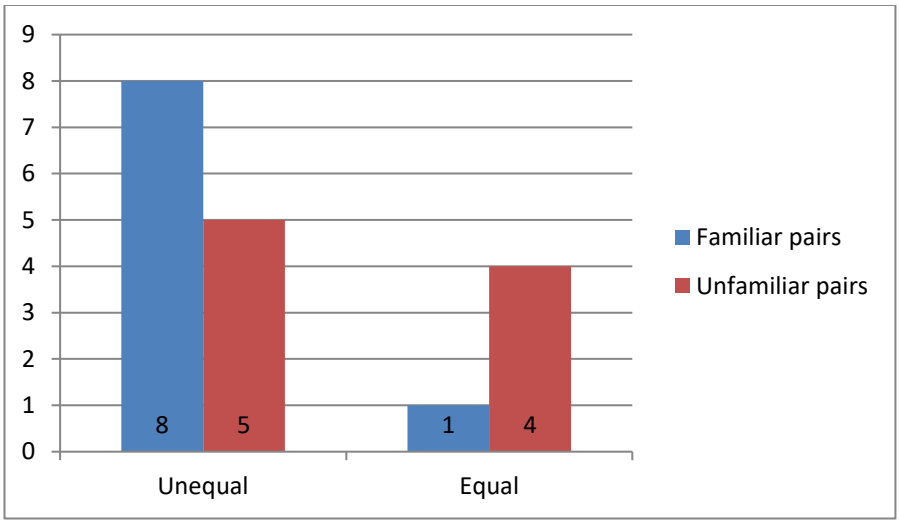


Figure 21. Distribution of turn dominance in familiar and unfamiliar pairs

As shown in Figure 21, while in eight conversations out of nine, one of the test-takers was more dominant than the other one, in only one conversation there was no dominance of any test-takers. In contrast, in the unfamiliar group, in five conversations out of nine, one of the test-takers was more dominant; however, in four conversations, there was no dominance at all. Hence, turn dominance was observed more frequently in familiar pairs in the present study.

All things considered, the results yielded that turn dominance seems to vary depending on the familiarity factor because familiar pairs seem more dominant than unfamiliar pairs in this paired speaking test.

Topic management in familiar and unfamiliar pairs. As stated before, in the present study, topic management was operationalized as topic initiation, topic extension, and topic shift and topic closure.

Topic initiation. To start with, topic initiation was analyzed in the scope of the task that the test-takers engaged in during the test discourse. To put it more simply, the test-takers were informed about who needs to start the conversation by means of the role cards which were assigned to them by the examiner. For this reason, while examining topic initiation in familiar and unfamiliar pairs, whether the topic was started by the test-taker that the task required, by the test-taker that the task did not require or by the examiner was taken into account.

For the samples of topic initiation, please see Excerpt 5 and 6 (Table 8 and 9).

Figure 22 shows the distribution of topic initiation in familiar and unfamiliar pairs (See Appendix K.5 and L.5 for the descriptive statistics of topic initiation in familiar and unfamiliar pairs).

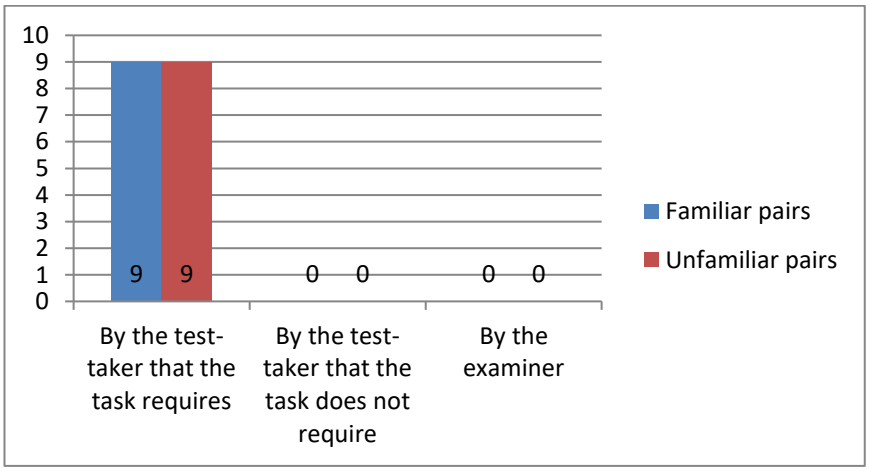


Figure 22. Distribution of topic initiation in familiar and unfamiliar pairs

As demonstrated in Figure 22, familiar pairs were able to start the topic when the task required them to in all nine conversations which were analyzed. Similarly, in unfamiliar pairs, the topic was initiated by the test-taker that the task required in all nine conversations again.

Overall, the findings indicated that familiarity does not influence the topic initiation of the test-takers because no matter they are paired with a familiar or an unfamiliar peer interlocutor; they can start the topic when it is required.

Topic shift. When it comes to topic shift in familiar and unfamiliar pairs, while analyzing it, whether test-takers shifted the topic and if yes, whether they were equal was considered. It is significant to note that, as stated before, because of the nature of the tasks given to the test-takers, in some conversations, only one test-taker was allowed to shift the topic during the test-discourse. For this reason, topic shift was not considered as the indicator of turn dominance as suggested by some researchers; rather it was regarded as only a part of topic management.

For an example of topic shift, please see Excerpt 7 (Table 10).

Figure 23 shows the distribution of topic shift in familiar and unfamiliar pairs (See Appendix K.6 and L.6 for the descriptive statistics of topic shift in familiar and unfamiliar pairs).

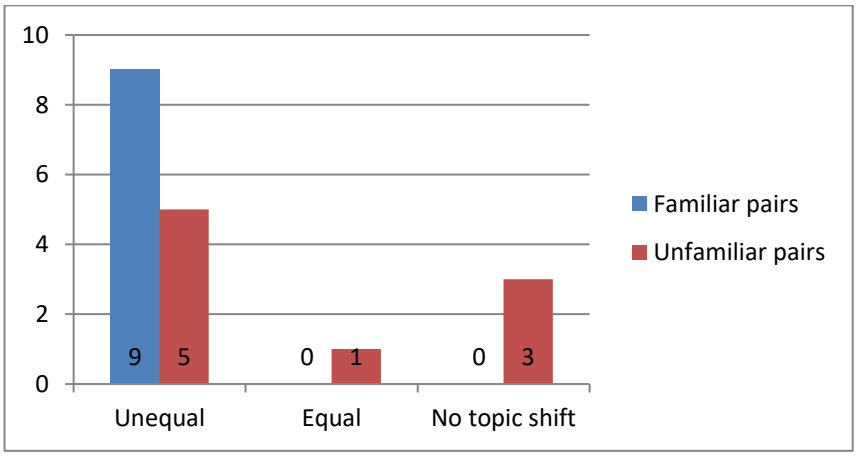


Figure 23. Distribution of topic shift in familiar and unfamiliar pairs

As demonstrated in Figure 23, in all nine conversations, the topic was shifted; however, merely one test-taker shifted it. Therefore, the test-takers were unequal in topic shift. However, in three conversations held by unfamiliar pairs out of nine, there was no topic shift during the test discourse. On the other hand, in the six conversations in which the topic was shifted by the test-takers, while both test-takers changed the topic, in only one conversation, the topic was shifted by one test-taker in the other five conversations.

All in all, it emerged from the findings that topic shift seems to differ depending on the familiarity factor. When two familiar test-takers are paired together, they can shift the topic even though it is unequal (i.e., one of them shifts more than the other one). On the other hand, unfamiliar pairs sometimes cannot shift the topic during their interaction with each other. However, when they shift the topic, they are also unequal like familiar pairs.

Topic extension. Another subcomponent of topic management is topic extension which is an integral part of it. While examining whether students' proficiency levels play a role on their topic extension, two perspectives were discussed: 1) who extended the topic, 2) to what extent the topic was extended. Nevertheless, while analyzing the effect of familiarity factor on topic extension, only the first one was taken into consideration for the reason that the main aim in familiarity factor was to learn about if the test-takers were equal or unequal in terms of the usage of interactional resources one of which is topic extension.

For the samples of topic extension, please see Excerpt 8 and 9 (Table 11 and 12).

Figure 24 illustrates the distribution of topic extension in familiar and unfamiliar pairs (See Appendix K.7 and L.7 for the descriptive statistics of topic extension in familiar and unfamiliar pairs).

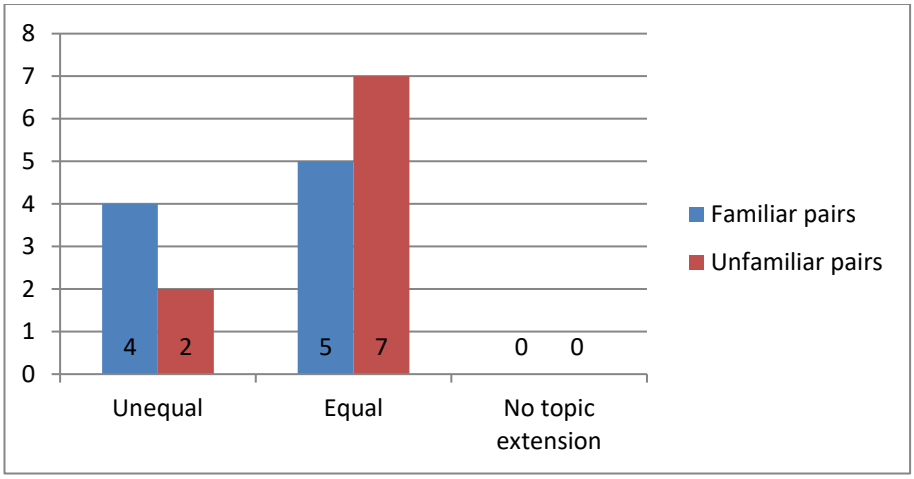


Figure 24. Distribution of topic extension in familiar and unfamiliar pairs

As shown in Figure 24, in five conversations out of nine, both test-takers attempted to extend the topic by asking questions to his/her peer interlocutor or by expressing further ideas about the topic discussed during the test discourse. Nonetheless, in four conversations merely one test-taker tried to develop the topic. When it comes to unfamiliar pairs, out of nine conversations which were analyzed for this study, in seven of them both test-takers developed the topic during the test-discourse whereas in two conversations, the topic was extended by only one test-taker.

On the whole, familiarity does not seem to play a role on the topic extension in paired speaking tests because both familiar and unfamiliar pairs can extend their topics during the test discourse, but it seems to affect the equality (i.e., who extends the topic) of test-takers. Unfamiliar pairs work more collaboratively as they both contribute to the talk than familiar pairs do.

Topic closure. The last constituent of topic management is topic closure which was also analyzed within the context of the present study. While analyzing topic closure, if there was topic closure during the test discourse was considered along with whether test-takers were equal in terms of topic closure.

For the examples of topic closure, please see Excerpt 12, 13 and 14 (Table 15, 16 and 17).

Figure 25 indicates the distribution of topic closure in familiar and unfamiliar pairs. (See Appendix K.8 and L.8 for the descriptive statistics of topic closure in familiar and unfamiliar pairs).

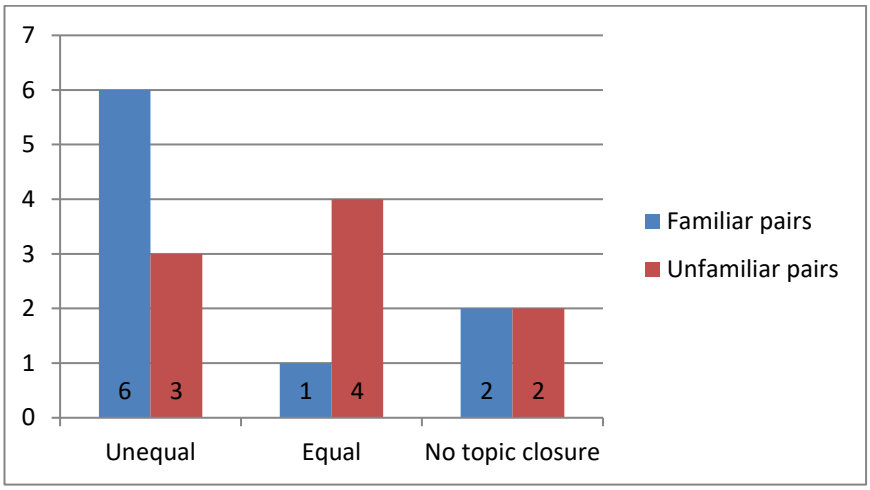


Figure 25. Distribution of topic closure in familiar and unfamiliar pairs

As it is obvious in Figure 25, in only two conversations out of nine, the peer interlocutors were not able to close the topic during their interaction. On the contrary, the topic was closed by the test-takers in the other seven conversations; however, while in six of them only one test-taker closed the topic and the other one did not, in merely one conversation, both test-takers were able to end the topic. As can be seen in Figure 25 again, similar to the results in familiar pairs, in two conversations out of nine, the topic was not closed by the test-takers in unfamiliar pairs. In contrast, in unfamiliar pairs, the number of the conversations in which the topic was closed by

both test-takers was more than the one in familiar pairs. While in four conversations, the topic was ended by both test-takers, in three of them, only one test-taker closed the topic and the other one did not.

Overall, the findings yielded that the topic seems to be closed by both familiar and unfamiliar pairs, but the equality of test-takers while closing the topic differs depending on the familiarity factor. To put it another way, in unfamiliar pairs, the topic is usually closed by both test-takers; however, in familiar pairs, only one test-taker usually closes the topic.

Repair in familiar and unfamiliar pairs. Although the main aim was to look into if the test-takers were equal or not, while analyzing the familiarity factor, the researcher did not focus on the main aim while analyzing repair in familiar and unfamiliar pairs. In this sense, considering that test-takers can repair the trouble when it is required and they do not resort to repair since it is not required or even though it is required, the categories were identified in that way.

For the samples of repair, please see Excerpt 15, 16 and 17 (18, 19 and 20).

Figure 26 illustrates the distribution of repair in familiar and unfamiliar pairs (See Appendix K.10 and L.10 for the descriptive statistics of repair in familiar and unfamiliar pairs).

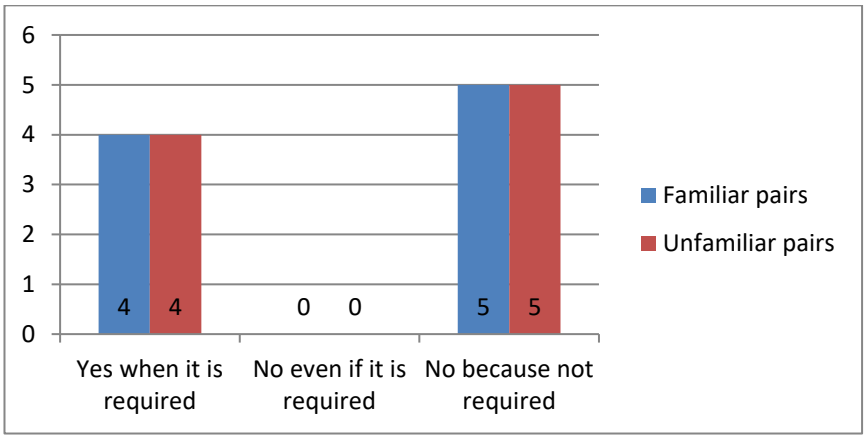


Figure 26. Distribution of repair in familiar and unfamiliar pairs

As indicated in Figure 26, in four conversations out of nine, the familiar test-takers were able to repair the trouble when it was required during the test discourse. On the other hand, in five conversations, they did not resort to repair as it was not required during their interaction with each other. Similar to the results in familiar pairs, while in four conversations out of nine, repair was applied when it was needed, in five conversations there was no repair because it was not required in unfamiliar pairs.

On the whole, the findings revealed that familiarity definitely does not play a role on repair in paired tests since in both familiar and unfamiliar pairs; the test-takers can repair the trouble when it is required. If they do not resort to repair, it results from that they do not need repair during their interaction.

Task management in familiar and unfamiliar pairs. The last interactional resource analyzed in familiar and unfamiliar pairs was task management. As mentioned earlier, while analyzing task management, whether test-takers help each other accomplish the task in order to meet the requirements of the task was taken into consideration. Moreover, whether there was task management during the test discourse, and if yes, whether they were equal while managing the task was analyzed.

For the samples of task management, please see Excerpt 18, 19 and 20 (Table 21, 22 and 23).

Figure 27 shows the distribution of task management in familiar and unfamiliar pairs (See Appendix K.9 and L.9 for the descriptive statistics of task management in familiar and unfamiliar pairs).

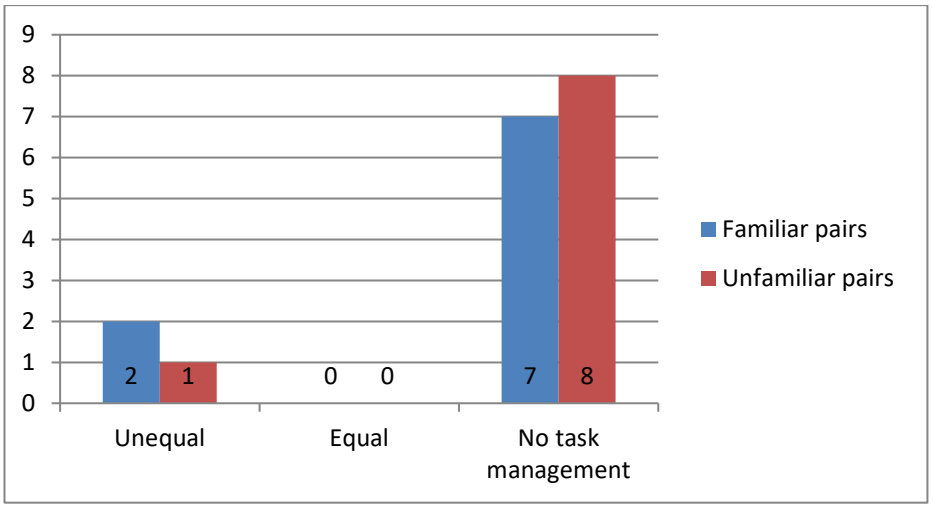


Figure 27. Distribution of task management in familiar and unfamiliar pairs

As illustrated in Figure 27, in seven conversations out of nine, there was no task management at all for the reason that the test-takers accomplished the task without any help. On the contrary, the task was managed in two conversations, but it was done by only one test-taker. Similarly, in eight conversations out of nine, the task was not managed by any test-takers, whereas in merely one conversation it was managed by one of the test-takers in unfamiliar pairs.

All in all, it was found out that task management seems not to differ depending on the familiarity factor inasmuch as both familiar and unfamiliar pairs do not manage the task during the test discourse. In other words, the test-takers do not help each other to stay on the task in both familiar and unfamiliar pairs.

Interactional patterns generated by familiar and unfamiliar pairs. Similar to the analysis of the proficiency factor, the second phase of the data analysis involved determining the interactional patterns such as collaborative, asymmetric, parallel and blend while examining the familiarity factor. As stated previously, in an attempt to identify the interactional patterns generated by familiar and unfamiliar pairs, the researcher benefitted from the interactional resources; turn-taking, topic management and task management found out in the first stage.

Interactional patterns in familiar pairs. Figure 28 illustrates the distribution of interactional patterns in familiar pairs. (See Appendix M for the descriptive statistics of interactional patterns in familiar pairs).

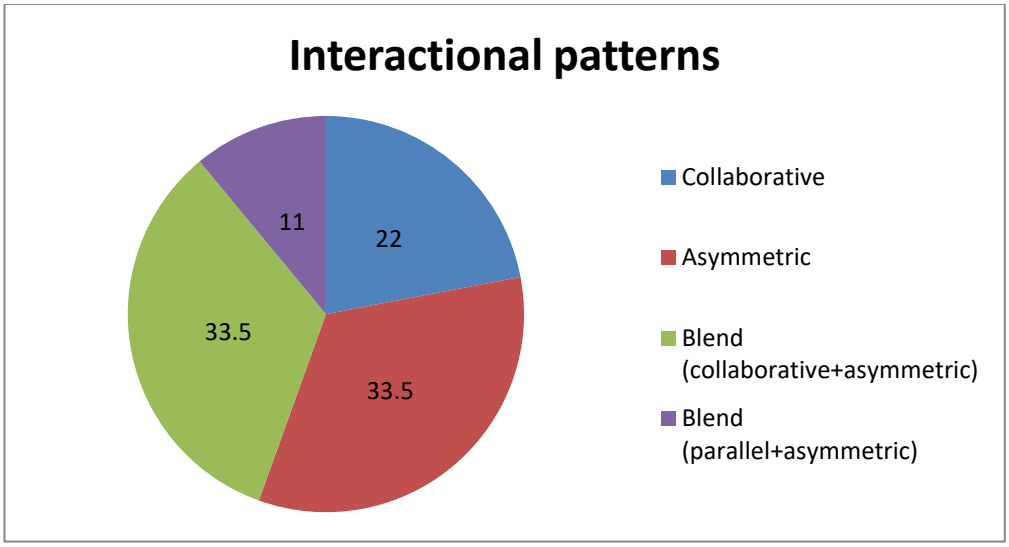


Figure 28. Distribution of interactional patterns in familiar pairs

As can be seen in Figure 28, the two mostly generated patterns in familiar pairs were asymmetric (33.5) and the blend of collaborative and asymmetric patterns (33.5). In the asymmetric pattern, the test-takers were unequal in terms of mutuality and equality between them. On the other hand, in the blend of collaborative and asymmetric patterns, the test-takers generated high mutuality and low equality, which means that they were able to manage the topic and task together and responded to the speaker as a listener, but one of them was not equal in terms of the quantity of talk s/he had. On the other hand, collaborative pattern itself was seen less frequently in familiar pairs with a percentage of 22 which corresponds to two conversations out of nine. Lastly, in 11 of the conversations, the pattern generated by familiar pairs was the blend of parallel and asymmetric patterns.

Interactional patterns in unfamiliar pairs. Figure 29 demonstrates the distribution of interactional patterns in unfamiliar pairs (See Appendix N for the descriptive statistics of interactional patterns in unfamiliar pairs).

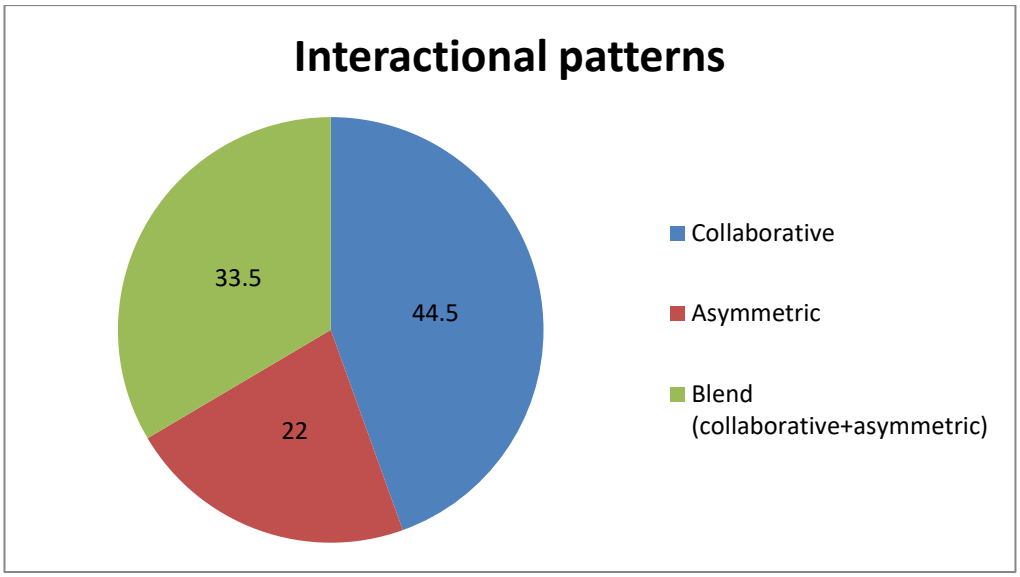


Figure 29. Distribution of interactional patterns in unfamiliar pairs

As shown in Figure 29, unlike the familiar pairs; the collaborative pattern was mostly generated one which includes high mutuality and high equality. In 44.5 of the conversations, unfamiliar pairs created collaborative pattern during the test discourse, which means that in most of the conversations, the test-takers extended the topic together with a balanced quantity of talk and they responded to one another when they were in the listener role. In addition, the blend of the collaborative and asymmetric patterns, which consists of high mutuality and low equality, was generated by the test-takers with a percentage of 33.5. Lastly, in 22 percent of the conversations, the pattern was asymmetric in which one of the test-takers was more dominant both in task and topic management and in turn-taking.

Table 25 presents a summary of the findings regarding the interactional patterns in familiar and unfamiliar groups.

Table 25

The Frequency of Interactional Patterns in Familiar and Unfamiliar Pairs

	Familiar pairs (%)	Unfamiliar pairs (%)
Collaborative	22	44.5
Asymmetric	33.5	22
Parallel	0	0
Blend (collaborative + asymmetric)	33.5	33.5
Blend (parallel + asymmetric)	11	0

Overall, when the interactional patterns that took place during the test discourse in familiar and unfamiliar pairs were analyzed, it emerged from the findings that they varied depending on the familiarity factor. While unfamiliar pairs mostly generated a collaborative pattern, this pattern was observed in familiar pairs more rarely. The two patterns mostly observed in familiar pairs were asymmetric and the blend of collaborative and asymmetric patterns. Similarly, the blend of collaborative and asymmetric patterns was found out in unfamiliar pairs at the same percentage with the one in familiar pairs. However, unfamiliar pairs generated the asymmetric pattern more rarely. Lastly, the blend of parallel and asymmetric patterns took place during one conversation held by familiar pairs.

Conclusion

In this chapter, the data gathered from the 50 videos recorded during the oral proficiency exam at the School of Foreign Languages of Bülent Ecevit University were analyzed descriptively and covered in two main sections. In the first section, whether the use of interactional resources and the interactional patterns vary when two same and two different proficiency level test-takers are paired was explained. In

the second section, whether the use of interactional resources and the interactional patterns differ when two familiar and unfamiliar test-takers are paired was presented. The next section will provide the findings in relation to the relevant literature along with the pedagogical implications, limitations of the study and suggestions for further research.

CHAPTER V: CONCLUSION

Introduction

The aim of this descriptive study was to investigate whether interactant related factors (i.e., familiarity and proficiency) play a role in EFL learners' a) use of interactional resources and b) the emergence of interactional patterns in paired speaking tests. In this respect, the present study addressed the following research questions:

1) How do the use of interactional resources and the emergence of interactional patterns vary in paired speaking tests with EFL test-takers

a) who have the same vs. different proficiency levels?

b) who are familiar vs. unfamiliar with each other?

This chapter involves four sections. In the first section, the findings of the present study will be discussed in light of the relevant literature. In the second section, the pedagogical implications will be presented. In the next section, the limitations of the study will be introduced, and in the last section, suggestions for further research will be provided.

Findings and Discussion

The Use of Interactional Resources in Paired Speaking Tests with EFL Test-takers Who Have the Same vs. Different Proficiency Levels

When the interactional resources were analyzed in the 50 videos in which test-takers were paired as high-high, low-low and low-high, the results indicated that while proficiency has a role on most of the interactional resources (i.e., turn speed, turn length, turn dominance, topic shift, topic extension, topic closure, repair and task management), only two of the resources (i.e., overlap and topic initiation) do not

differ depending on the proficiency factor.

Table 26 below provides the use of interactional resources in three different groups (i.e., high-high, low-low and low-high).

Table 26

The Use of Interactional Resources Depending on the Proficiency Factor

Interactional resources		**Pairing different proficiency level test-takers plays a role or not?**	**High-High**	**Low-Low**	**Low-High**
Turn-taking	Overlap	No	There is overlap.	No overlap.	No overlap.
	Turn speed	Yes	Both test-takers usually speak at the same speed.		High level test-taker is faster.
	Turn length	Yes	Both test-takers usually have the same turn length.		High level test-taker takes more and/or longer turns.
	Turn dominance	Yes	There is usually no dominance.		High level test-taker is more dominant.
Topic management	Topic initiation	No	They can initiate the topic when the task requires them to.		
	Topic extension	Yes	They usually extend the topic collaboratively, so topic extension is satisfactory.	They usually extend the topic individually, so topic extension is limited.	High level test-taker usually extends the topic, so topic extension is limited.
	Topic shift	Yes	There is topic shift.	No topic shift.	Topic shift is performed only by the high level test-taker.
	Topic closure	Yes	They can close the topic.	No topic closure.	No topic closure.
	Repair	Yes	There is repair when it is needed.	There are instances of lack of repair even though it is needed.	Repair is performed by the high level test-taker.
	Task management	Yes	No task management.		High level test-taker manages the task.

As can be seen in Table 26, while overlap and topic initiation do not differ, the other interactional resources (i.e., turn speed, turn length, turn dominance, topic shift, topic extension, topic closure, task management and repair) vary depending on the proficiency factor. The next sections will provide information about the role of proficiency factor in the use of interactional resources and interactional patterns in detail.

Turn-taking. According to Feldstein and Welkowitz (1987) during a talk speakers need to talk one by one and they should be aware of when they will end their own turn and allow the next speaker to speak, which refers to avoiding overlaps while taking turns. Turn-taking management consists of turn length, turn speed and turn domination (Ducasse & Brown, 2009; Ducasse, 2010; Watanabe & Swain, 2007) and speakers are expected to have a no-overlap conversation during their talk. When these four subcomponents (i.e., overlap, turn length, turn speed and turn dominance) of turn-taking was examined in this study, it emerged from the findings that when test-takers are paired with another test-taker who is at the same proficiency level, their turn speed is usually the same, they usually have the same number and/or quantity of turns, and there is no dominance during the test discourse. However, when high levels are matched with low level test-takers, high level ones are mostly faster than the low level one while taking turns because low level test-takers speak with many pauses during the test discourse which is also suggested by Galaczi (2013). Furthermore, the fact that low level test-takers are slower while responding to high level test-taker leads the high level speakers to taking more turns in the oral exam because high level test-takers support the speaker with minimal responses and agreements when they are in the listener role; however, low levels' listener support is limited as also suggested by Gan (2010) and Galaczi (2013). In addition, high levels

feel the need to encourage the low level speakers to speak by asking further questions or to explain the task during the waiting period, which ends up with longer and/or more turns. The fact that high level test-takers take longer and/or more turns than the low levels do in the low-high group is also an indicator of turn dominance. Since high level test-takers' turn length is longer and they stay on the floor more, they are more dominant than the low level ones in turn-taking. Considering these findings, it can be argued that turn-taking management which includes turn speed, turn length and turn dominance requires a speaker to have a high level of interactional competence. Since high level test-takers have a better level of interactional competence, they are better at turn-taking management, which is also suggested by Galaczi (2013). With this in mind, it can be concluded that when two different proficiency level test-takers are paired together, low level test-takers can be at the disadvantageous side. They may feel anxious since they cannot respond to their high level partners at the same speed and they cannot take as long turns as their partners do.

On the other hand, as the results indicated overlap is not observed when different proficiency level test-takers are paired together because overlaps occurred only in high-high pairs, which is in line with Gan's (2010) and Galaczi's (2013) studies. In Gan's (2010) study, the overlaps were regarded as the indicator of other-initiated topic extension. In other words, since the high level test-takers are willing to engage in not only their own, but also their partners' topics, which will be discussed in detail later on, these dyads cause overlaps. Moreover, high level test-takers usually compete for the floor to take control of the interaction, which can be another reason of overlaps. However, when a low level test-taker comes into play, no-overlap conversation is held by the test-takers even though they are paired with high levels

because they do not struggle for the floor and they usually do not contribute to their partners' topics.

Topic management. Another interactional resource analyzed in the scope of this study is topic management and under this resource, topic initiation, topic extension, topic shift and topic closure were examined because having a smooth conversation requires speakers to know how to initiate the topic, how to extend and shift it and when they need to close it (Dimitrova-Galaczi, 2004; Ducasse & Brown, 2009; Gan, 2010; May, 2011b). When the subcomponents (i.e., topic initiation, topic extension, topic shift and topic closure) of topic management were analyzed, it emerged from the findings that topic extension, topic shift and topic closure differ depending on the proficiency factor, while topic initiation does not.

When the test-takers who have the same proficiency levels are examined, the findings yielded that the test-takers paired as high-high usually extend the topic together by engaging in each other's topics as well, which was also found out by Gan (2010) and Galaczi (2013). The higher the level of the students is, the more they have other-initiated topic moves. The findings also revealed that high-high pairs can shift the topic in the exam and close the topic without the intervention of the examiner. In contrast, low-low pairs cannot extend the topic collaboratively; rather, they engage in only their own topics and work individually while developing the topic, so their topic extension is highly limited during a paired test discourse. What is more, they cannot shift the topic during the test discourse and can never close the topic themselves. The conversation is always ended by the examiner.

On the other hand, when low levels are paired with high levels, their topic extension is limited as in the low-low group due to the restricted contribution of low level test-takers to topic extension. Moreover, in low-high pairs, the topic shift is

performed by high level test-takers because topic shift is usually associated with the dominance of one speaker and considering that high level test-takers are more dominant when they are matched with low levels, they mostly shift the topic. Furthermore, low levels' lack of ability to close the topic also influences the high level test-takers in low-high pairs because the conversations held by low-high dyads are also ended by the examiner, not by the test-takers. The underlying reason for this perhaps rests on that because the low level test-takers talk and contribute less; the high level ones make more effort in low-high pairs than they would do in high-high pairs. Because of high levels' concern about managing the topic and task and keeping the conversation going; they may be missing when they need to close the topic. Additionally, low level test-takers' limited topic management can be linked to that topic management requires a high level of interactional competence and the concept of interactional competence is considered to be a joint work of two speakers due to its co-constructed nature (Kramsch, 1986; He and Young, 1998). Interactional competence entails the interlocutors to collaborate and contribute to the talk together. They should be able to negotiate meanings and create intersubjectivity between each other (Young, 2008). Therefore, it can be claimed that pairing two different proficiency level test-takers can be disadvantageous from the high levels' perspective in terms of topic management because low level test-takers cannot contribute to the talk much since they lack the interactional competence.

On the other hand, it emerged from the findings that topic initiation is achieved by the test-takers in all three groups (i.e., high-high, low-low and low-high) even though there were a few instances of no-topic-initiation in the low-low group. This result is not surprising when high-high and low-high groups are considered because the high level test-taker understands the task well and knows how to initiate

the topic to maintain the conversation. However, the fact that test-takers paired as low-low can also initiate the topic may be resulting from that topic initiation was controlled by the task in the present study since the task assigned who needs to start the conversation to the test-takers. Therefore, the test-takers may have felt the obligation to initiate the topic in a way. Moreover, this result indicates that low level test-takers do not have difficulty in starting the conversation, but they get in trouble in maintaining it because they cannot extend the topic and shift it.

Task management. The findings indicate that there is no task management when two same proficiency level test-takers are paired together. In high-high pairs, the help of the partner is not needed while managing the task because they can accomplish the task without any help, whereas in low-low pairs, task management is required, but they cannot help each other because of their low proficiency levels. On the other side, when low level test-takers are paired with a high level peer interlocutor, the high level ones usually manage the task in order to meet its requirements. For this reason, it can be concluded that pairing two different proficiency level test-takers can be advantageous from low levels' perspective. However, it can be disadvantageous for the high level ones because they make extra effort to accomplish the task correctly as they not only manage their part in the task, but also help their low level peer interlocutors.

Repair. The results regarding repair indicated that high-high pairs can repair the trouble whenever it is needed, but there is usually no communication breakdown, so they do not resort to repair. On the other hand, low-low pairs usually need repair; however, they sometimes cannot repair the trouble even though it is required during the talk. What is more, when a high level and a low level test-taker matched together, repair is needed more and mostly the high levels repair the trouble most probably

because the low level test-takers respond to the high levels late, they think that their low level partners do not understand what has been uttered. Furthermore, applying repair strategies requires one to have a good command of the communication skills which is related to interactional competence because repair strategies are employed when there are communication problems (Faerch & Kasper, 1983). Since the low level test-takers do not have a high level of interactional competence, they cannot use the repair strategies effectively. For this reason, it can be claimed that pairing two different proficiency level students can be disadvantageous from the high level test-takers' side because when the high level test-taker is paired with a low level test-taker, his/her low level peer interlocutor may hinder the flow of the conversation since the low level ones sometimes cannot repair the trouble and this may cause communication breakdowns.

The Interactional Patterns in Paired Speaking Tests with EFL Test-takers Who Have the Same vs. Different Proficiency Levels

According to Galaczi (2008) interactional patterns (i.e., collaborative, asymmetric, parallel and blend) are generated depending on how interactional resources are employed by the test-takers in a paired speaking test discourse. When the three groups paired as high-high, low-low and low-high are examined, findings of this study suggest that pairing two different proficiency level test-takers seem disadvantageous for both the low levels and the high levels since low-high pairs create an asymmetric pattern. It can be argued that high level test-takers have a greater level of responsibility when they are paired with a low level one and they feel the need to deal with both their low level partners' topic and their own, which prevents them from employing some of the interactional resources (e.g., topic closure) effectively. Moreover, low level test-takers always feel the dominance of

high levels, which may cause them to feel more anxious during the exam. On the other hand, when two same proficiency level test-takers are paired together, the high-high pairs seem the most advantageous group among the three groups (i.e., high-high, low-low and low-high) in a paired speaking test as this group creates a collaborative pattern during the test discourse. Inasmuch as both peer interlocutors are good at turn-taking, topic management and task management owing to their high level of interactional competence, they easily maintain the conversation together.

When peer interlocutors are paired as high-high, they mostly generate a collaborative pattern which involves high equality and high mutuality. High equality is observed in high-high pairs because these test-takers' quantity of talk is almost the same which refers to turn length and they respond to each other almost at the same speed which refers to turn speed. Moreover, since high level test-takers have interactional competence, they can engage in other-initiated topics, which is associated with high mutuality. In other words, they extend not only their own topics, but also their partners' topics and they support the speaker when they are in the listener role during the test discourse. For these reasons, the most frequent pattern generated in high-high pairs is the collaborative one. This finding shows parallelism with Galaczi's (2008) study because she also found that the test-takers who created a collaborative pattern during the test discourse got the highest scores in the speaking exam that she analyzed. Additionally, the blend of collaborative and asymmetric patterns is also observed when two high level test-takers are matched together. In the conversations in which this pattern is generated, the test-takers display high mutuality and low equality. While they contribute to each other's topics by working collaboratively and support each other when they are in the listener role, they differ

in their turn length and in the extent to which they use the minimal responses and the agreements to support the speaker.

On the other hand, when two test-takers are paired as low-low, they mostly create a parallel pattern which includes high equality and low mutuality. In these dyads, high equality occurs because their turn length and turn speed are very similar and there is no turn dominance. However, low-low pairs do not engage in other-initiated topics, but they usually extend self-initiated topics during the test discourse, which can be referred to as low mutuality. To put it another way, they work individually and do not contribute to their partners' topics. What is more, their listener support is highly limited during the interaction as they use the minimal responses and agreements to a small extent. This finding is concurrent with Galaczi's (2008) study again since it emerged from her findings that the test-takers who generated a parallel pattern got the lowest scores in the exam. Different from Galczi's (2008) study, however, in this study, the blend of parallel and asymmetric patterns in which there is low mutuality and low equality was also found out in low-low pairs. Seeing that in some conversations, one of the test-takers is more dominant in terms of his/her quantity of talk and the extent to which s/he supports the speaker, but both test-takers still work individually without contributing to each other's topics, it can be claimed that low-low pairs create a blend of parallel and asymmetric patterns as well as a parallel pattern. Although personality factor was not taken into consideration in this study, test-takers' imbalance in their quantity of talk and degree of listener support can be attributed to shyness of some test-takers because in both Berry's (2004) and Bonk and Van Moere's (2004) studies, it was found out that outgoing test-takers are more dominant than the shy ones during the test discourse.

Moreover, when a low level test-taker is matched with a high level one, the most frequent pattern is the asymmetric one which consists of moderate mutuality and low equality. The reason why asymmetric pattern is the most frequent one in low-high pairs is that in those pairs, there is inequality in turn speed, turn length and degree of listener support since high level test-taker is more dominant while employing these resources. Moreover, moderate mutuality results from that the topic is mostly extended by the high level test-taker and the low level ones do not contribute to the talk much. Therefore, they cannot create a shared meaning during their interaction (Dimitrova-Galaczi, 2004). Different from high-high and low-low pairs, in the low-high group, task management (i.e., high level test-takers help the low levels accomplish the task correctly) is observed too. While managing the task, high level test-takers engage in their partners' topics as well to help them meet the requirements of the task, which can also be associated with moderate mutuality since there is no contribution of low levels to the conversation at this point.

Discussion of the Findings Regarding Proficiency as a Factor in Paired Speaking Tests

Although paired speaking tests are regarded as more advantageous than individual tests since they allow more interaction during the test discourse, the implementation of paired tests is disputable due to test-taker characteristics such as proficiency, familiarity, personality and gender (McNamara, 1997; Swain, 2001; Young & He, 1998). Given that proficiency level is a factor which may play a role on the interaction between two test-takers in paired speaking tests, in this study, the first aspect focused on in the research question was whether the use of interactional resources and interactional patterns vary when two same and two different proficiency level test-takers are paired together. The findings of this study suggest

that pairing two different proficiency level test-takers seems disadvantageous in terms of topic management, task management and repair for the high proficiency level students as low level test-takers' contribution to the topic and task is quite limited. Additionally, low levels have difficulty in repairing the trouble when needed, which may cause communication breakdowns, during the test discourse. Therefore, high levels usually feel the need to take the responsibility of their low level partners' as well as their own and need to put extra effort to manage the topic and task during the test discourse.

On the other hand, low level test-takers are advantageous in terms of topic management, task management and repair when they are paired with a high proficiency level partner since their high level peer interlocutors can manage not only their own parts, but also their low level partners' parts during the exam. Moreover, high levels can easily repair the trouble when it is required, which low level test-takers mostly need. However, pairing low levels with high level ones can be disadvantageous for the low level test-takers in terms of turn-taking as they cannot speak as fast as their high level peer interlocutors and they cannot take as long and/or many turns as their high level partners do. High level test-takers support their low level peer interlocutors with minimal responses and agreements and ask further questions to encourage the low levels to speak, whereas low levels cannot. For this reason, they take more and/ or longer turns, which is also an indicator of dominance in turn-taking. Hence, the fact that high level test-takers are more dominant in turn-taking may create a pressure on low level test-takers and they may feel more anxious to speak because they speak with many pauses.

When it comes to interactional patterns generated by the test-takers, it can be claimed that interactional patterns differ depending on the pairing system in paired

speaking tests. While high-high pairs create a collaborative pattern, low-low pairs generate a parallel and the blend of parallel and asymmetric patterns. On the other hand, when two different proficiency level test-takers are matched together, they create an asymmetric pattern. Among the three groups (i.e., low-low, high-high and low-high), the most favorable one is pairing two high level test-takers together. Since both of them have interactional competence and good at topic management, turn-taking and task management, they can work collaboratively. The only handicap in pairing high-high test-takers is that overlaps occur only in this group. Since they are both willing to stay on the floor, they cause more overlaps. On the other hand, when high levels are paired with low levels, high level test-takers naturally become more dominant during the test discourse, which turns out to be the emergence of the asymmetric pattern. The fact that there is an asymmetric pattern in low-high dyads is definitely disadvantageous for the high level test-takers because they have to make extra effort to maintain the conversation as low level test-takers' contribution to the task and topic is limited. On the other hand, this can be regarded as an advantage for the low levels in terms of topic and task management in particular as they can produce more talk with the help of high levels and maintain the conversation better. Nevertheless, considering turn-taking management, the dominance of the high level test-takers may cause anxiety on the low level test-takers. Since they feel less competent in front of their more competent peer interlocutors, they might be taking less and/or shorter turns during the test discourse.

According to Eygud and Glover (2001) in a paired speaking test, the speakers have a more balanced interaction in terms of their quantity of talk and topic initiation and extension, which is proposed as an advantage of paired tests. Moreover, it is claimed that the speakers have less communicative anxiety and stress as they are

tested with their peers (Ikeda, 1998; Norton, 2005; Saville & Hargreaves, 1999). Paired speaking tests are also regarded as having stronger construct validity than the individual interviews do since paired tests enable authentic conversation with equal turn taking opportunities (Green 2014; Okada, 2010; Van Lier 1989). However, these results suggest that paired tests do not have an advantage when two different proficiency level test-takers are paired together. The resources such as turn-taking, topic management, task management and repair all require speakers to have a high level of interactional competence and communication skills. However, when a low-high pair is matched together, they do not have balanced turn-taking opportunities as high level ones usually take longer and/or more turns. Furthermore, since low level test-takers lack the interactional competence, they cannot work collaboratively with their high level partners to manage the task and topic. For these reasons, it can be concluded that pairing two different proficiency level test-takers destroys the advantages of paired speaking tests as this pairing system influences the interaction negatively.

The Use of Interactional Resources in Paired Speaking Tests with EFL Test-takers Who Are Familiar vs. Unfamiliar with Each Other

When the interactional resources were analyzed in the 18 videos in which nine pairs were familiar and the other nine dyads were unfamiliar with each other, the findings revealed that while familiarity does not play a role on some of the interactional resources (i.e., turn speed, topic initiation, repair and task management), the use of other resources (i.e., overlap, turn length, topic shift, turn dominance, topic extension and topic closure) differ depending on the familiarity factor. Hence, it can be argued that familiarity does not have a role on some certain interactional resources in paired speaking tests.

Table 27 below presents the interactional resources which differ depending on the familiarity factor.

Table 27

The Use of Interactional Resources Depending on the Familiarity Factor

Interactional resources		Familiarity with the other test-taker plays a role or not?	Familiar	Unfamiliar
Turn-taking	Overlap	Yes	Although lack of overlap is more frequent in both familiar and unfamiliar pairs, more overlap is observed in familiar pairs.	
	Turn speed	No	They speak at the same speed.	
	Turn length	Yes	Familiar test-takers are usually more unequal in terms of turn length.	
	Turn dominance	Yes	There is more dominance of one test-taker in familiar pairs.	
Topic management	Topic initiation	No	They both initiate the topic when the task requires them to.	
	Topic shift	Yes	There is always topic shift in familiar pairs, but one test-taker shifts the topic, so they are unequal. There are the instances of no topic shift in unfamiliar pairs. When there is topic shift, they are also unequal.	
	Topic extension	Yes	There is always topic extension, but unfamiliar pairs work more collaboratively.	
	Topic closure	Yes	There is usually topic closure, but familiar pairs are usually more unequal.	
	Task management	No	No task management.	
	Repair	No	They repair the trouble when it is needed.	

As is clear in Table 27, while turn speed, topic initiation, repair and task management differ, the other resources (i.e., overlap, turn length, topic shift, turn dominance, topic extension and topic closure) do not vary depending on the familiarity factor. The following sections will present the role of familiarity factor in the use of interactional resources and interactional patterns in detail.

Turn-taking. Turn-taking consists of the subcomponents; overlap, turn speed, turn length and turn dominance. When these four subcomponents were analyzed, it emerged from the findings that both familiar and unfamiliar pairs' turn speed is the same. Thus, it can be claimed that test-takers' turn speed does not differ depending on the familiarity factor. On the other hand, lack of overlap is more frequent in both familiar and unfamiliar pairs; however, when they are compared to one another, familiar pairs have more overlaps. Furthermore, in familiar pairs, one test-taker takes longer and/or more turns than the other one, whereas in unfamiliar pairs, their turn length is usually equal. Accordingly, in familiar pairs, one test-taker is more dominant than the other one because turn dominance is basically determined by looking at the turn length of the speakers and their listener support (i.e., using minimal responses and agreements in the listener role). The underlying reason of the fact that familiar pairs have more overlaps, they are unequal in their number and/or quantity of talk, and there is dominance of one speaker in familiar pairs can be that when test-takers are tested with interlocutors with whom they are familiar, they feel more comfortable (O'Sullivan, 2002; Ying, 2009). Another reason may be that since they feel less anxious during the test discourse, they do not hesitate to take risks and compete for the floor, which ends up with overlaps and longer and/or more turns. Additionally, the result which suggests that one test-taker is more dominant than the other one in familiar pairs can be related to the case in individual interviews as

Okada (2010) proposes. Since the examiner manages the conversation mostly, the test-taker feels the dominance of the examiner during the test discourse in individual interviews and this situation leads to the generation of the asymmetric pattern in the exam. Because examiners are regarded as proficient user of the language by the test-takers, they may not feel comfortable enough to speak during the test discourse as test-takers perceive themselves as less competent. In light of Okada's (2010) suggestion, it can be argued that the reason why there is dominance in the conversations held by familiar pairs can be that because they are acquainted with each other's proficiency levels too, the peer interlocutors may have prejudices about their partners. While high level test-takers feel more confident while speaking, the low level ones may hesitate to speak with a high level partner, which turns out to be the dominance of high levels. Therefore, it can be concluded that pairing two familiar test-takers can be disadvantageous.

Topic management. As to the topic management, it includes the subcomponents; topic initiation, topic extension, topic shift and topic closure. When the subcomponents were analyzed, the findings yielded that unfamiliar pairs have more other-initiated topic moves. To put it another way, unfamiliar pairs extend their topics better and contribute to each other's topics more than the familiar pairs do. In the same vein to the result regarding topic extension, topic shift differs depending on the familiarity factor since in the conversations held by unfamiliar pairs there are the instances of no topic shift, but in familiar pairs there is always topic shift. However, in both familiar and unfamiliar dyads, when there is topic shift, only one test-taker shifts the topic not both of them. Similarly, when topic closure was examined, it emerged from the findings that there is topic closure in most of the conversations held by both familiar and unfamiliar pairs. Nonetheless, while in familiar pairs, only

one test-taker attempts to close the topic, in unfamiliar pairs, both test-takers close the topic. These results are unexpected in light of the suggestion by O'Sullivan (2002) that test-takers work better when they are paired with a peer interlocutor with whom they are familiar. Inasmuch as the proficiency factor was not controlled while examining the familiarity factor and most of the familiar pairs were matched as low-high in this study, the underlying reason of these results can be attributed to the familiarity to the proficiency of the partner again. Since familiar partners are acquainted with their proficiency levels, they may not be making much effort to help their partners maintain the conversation because they think that their partners are at low proficiency levels and they cannot accomplish the task in any way so they prefer to work individually. When this result is taken into consideration, pairing two unfamiliar partners can be more advantageous. Since the unfamiliar pairs do not have any idea about their partners' proficiency levels, they do not have any prejudices.

On the other hand, as the results indicate, familiarity does not play a role on topic initiation because both familiar and unfamiliar pairs can initiate the topic when the task requires them to. This result can be attributed to the nature of tasks in the study as in the proficiency factor because the tasks assigned the test-takers who will start the conversation. Furthermore, this result indicates that both familiar and unfamiliar pairs do not have difficulty in initiating the topic, which means that topic initiation does not differ depending on the familiarity factor.

Task management. When it comes to the task management, the findings yielded that in both familiar and unfamiliar pairs, no task management was observed. In other words, none of the test-takers helped each other to achieve the task during the test discourse. Thus, it can be claimed that task management does not vary depending on the familiarity factor. The reason for this perhaps rests on that both

familiar and unfamiliar pairs do not feel the need to help their partners accomplish the task correctly maybe because they can achieve the task without any help thanks to their high proficiency levels or because they cannot help each other because of their low proficiency levels even though it is needed.

Repair. When repair was examined, the findings yielded that familiarity does not play a role on repair since both familiar and unfamiliar dyads can repair the trouble whenever it is required. As stated previously, while analyzing the familiarity factor, the test-takers' proficiency level was not controlled. To put it differently, they were not selected from a particular group (i.e., low-low, high-high, and low-high), but they were selected randomly. Repair strategies require the test-takers to have a good command of communication skills (Faerch &Kasper, 1983). Considering that all familiar and unfamiliar pairs managed to use repair strategies in the present study, it can be concluded that repair is related to the proficiency of the test-takers and it does not differ depending on the familiarity factor.

The Interactional Patterns in Paired Speaking Tests with EFL Test-takers Who Are Familiar vs. Unfamiliar with Each Other

In the research question, another aspect focused on was the role of familiarity on the emergence of interactional patterns (i.e., collaborative, parallel, asymmetric and blend). In this sense, the results of this study suggest that pairing two unfamiliar pairs seem more advantageous since they usually create a collaborative pattern during the test discourse while the familiar ones usually generate an asymmetric pattern.

When test-takers are familiar with each other, they mostly create an asymmetric pattern or the blend of asymmetric and collaborative patterns. The existence of asymmetric pattern in most of the conversations results from that when

test-takers are familiar with each other; one test-taker is usually more dominant than the other one during the test discourse. More specifically, in familiar dyads, one test-taker takes longer and/or more turns and supports the speaker more with minimal responses and agreements when s/he is in the listener role. This result contradicts the findings of Porter (1991a) since he found out no acquaintanceship effect when 13 Arab learners were tested with familiar and unfamiliar examiners. However, Porter's (1991a) finding may be resulting from the fact that he investigated the role of familiarity in an individual test. In an individual test discourse which is carried out in the presence of an examiner and a test-taker, there is often the asymmetric pattern because both speakers do not have equal rights while speaking (Green, 2014; Van Lier, 1989). Nevertheless, in the present study the familiarity factor was explored in a paired test discourse and two test-takers have equal rights in such a discourse. Within this context, as familiar partners are acquainted with their proficiency levels, if low levels are aware of that they are matched with a high level test-taker they may perceive it in the same way with the one in an individual test discourse and it may cause the creation of the asymmetric pattern in familiar pairs. Moreover, high level test-takers may think that they are better than their partners, so they may struggle for the floor, which turns out to be the dominance of the high level test-taker. However, when they are unfamiliar with each other, the least frequent pattern created by the test-takers is the asymmetric one and they mostly generate a collaborative pattern. To put it differently, they usually work collaboratively while extending the topic and they are usually more equal in turn-taking. The reason why collaborative pattern was the mostly observed one in unfamiliar pairs can be that because they do not have any idea about their partners' proficiency levels, they feel the need to contribute to the talk in a way.

Discussion of the Findings Regarding Familiarity as a Factor in Paired Speaking Tests

Considering that familiarity with the other test-taker may be one of the factors which has role on the interaction in a paired speaking test discourse, another aspect focused on in the research question was the role of familiarity in the use of interactional resources and interactional patterns. In this regard, the findings of this study suggest that pairing two unfamiliar test-takers seem more advantageous in a paired speaking test because when they are paired together, they work more collaboratively than the familiar pairs and there is usually no dominance in the unfamiliar group. Pairing two unfamiliar test-takers is more favorable especially in terms of turn-taking and topic management as in unfamiliar dyads both test-takers contribute to the task and topic. Moreover, they usually take equal turns than the familiar pairs do. On the other hand, in familiar ones, one test-taker takes longer and/or more turns so there is the dominance of one speaker during the test discourse. For this reason, while unfamiliar pairs usually create a collaborative pattern which is the favorable one, familiar pairs usually generate an asymmetric pattern or the blend of collaborative and asymmetric patterns. All of these results can be attributed to the acquaintanceship with the proficiency level of the partner in familiar pairs. Since the familiar pairs are acquainted with their proficiency levels as well, they may have prejudices about their partners in the exam. While high level test-takers feel more confident and are more eager to take turns, the low levels may speak with hesitations when they know their partners' proficiency levels. What is more, while managing the topic and task, high level test-takers may think that their low level peer interlocutor cannot accomplish the task correctly even if they help their partner, which leads the

high level test-takers to working individually when they are paired with a low level partner. As a consequence, the high levels become more dominant in familiar pairs.

Paired speaking tests are claimed to be more authentic as they allow more balanced turn-taking opportunities to the test-takers (Eygud &Glover, 2001). In addition, they are regarded as more advantageous than the individual tests since it enables the test-takers to initiate and extend the topic equally. However, when two familiar pairs are matched together, the test-takers are acquainted with their partners' proficiency levels. For this reason, the familiar test-takers experience the case in an individual test discourse. Because the low level test-taker feels the dominance of the high level one, they become more passive compared to the high levels. Accordingly, these pairs create an asymmetric pattern as in an individual test.

Pedagogical Implications

The findings of the present study have some pedagogical implications for assessment. Since the main aim of the study was to investigate the role of proficiency and familiarity factors on interaction in paired speaking tests, it bears some possible implications for test administrators regarding how paired speaking tests should be implemented. The results of this study suggest that pairing two different proficiency level test-takers influence the high level test-takers in terms of topic management, task management and repair in a negative way and the low levels are affected negatively in terms of turn-taking management. Therefore, test administrators should avoid pairing two different proficiency level test-takers together. Considering the findings regarding interactional patterns, it can be suggested that the best pairing system is matching two high level test-takers together because they create a collaborative pattern owing to their higher level of interactional competence. Since both of them are good at employing the interactional resources, they both contribute

to the task and topic and they usually take equal turns. Hence, in order to achieve what paired speaking tests offer, high-high test-takers should be paired together. Moreover, pairing two familiar test-takers turns out to be the dominance of one peer interlocutor and self-initiated topic extension, and they usually generate an asymmetric pattern during the test discourse. Therefore, the results of the present study suggest that in order to make use of the advantages of paired speaking tests, it might be better to match two unfamiliar pairs.

Limitations of the Study

There are several limitations of this descriptive study suggesting that the findings should be interpreted cautiously. First of all, although the number of the video recordings analyzed to investigate the proficiency factor can be regarded as sufficient, the number of the videos used for exploring the role of familiarity factor was insufficient due to that most of the test-takers were unfamiliar with each other. In the study, 50 pairs selected for the proficiency group were all asked whether their partners were their friends or classmates and only nine pairs were familiar with each other. Therefore, in order to investigate the role of familiarity, only the videos of nine familiar and nine unfamiliar (18 in total) pairs were analyzed. For this reason, it may not be possible to generalize the findings obtained for the familiarity factor, in particular.

Another limitation is that in order to learn about test-takers' familiarity with each other they were asked if they were friends or classmates, but their degree of familiarity with each other (i.e., close friends, acquaintances, etc.) was not taken into consideration. Since the extent to which test-takers are familiar with each other may change the results, the findings regarding the familiarity factor should be interpreted cautiously.

Moreover, since this study is based on the video recordings retrieved from only one institution and the results may vary depending on divergent settings, different teaching approaches or the proficiency level of the test-takers accepted as low and high in a particular institution, it may not be possible to generalize the findings of the present study.

Suggestions for Further Research

In the present study, while analyzing the familiarity factor, proficiency level of the test-takers was not taken into consideration. In other words, while 12 familiar and unfamiliar pairs belonged to the high-high group, six of them belonged to the low-high group. However, they were not compared to one another on the basis of the group they belonged to, but they were compared randomly. Thus, in further studies, familiarity factor can be investigated by selecting the familiar pairs from a particular proficiency group and comparing them in themselves (e.g., only low-high group).

In addition, in this study, test-takers' degree of familiarity with each other was not taken into account. In further studies, their degree of familiarity with each other can be explored with a questionnaire first and the test-takers who have the same level of familiarity with their partners can be selected as participants of the study so as to gain a better understanding of the role of familiarity factor on interaction in paired tests.

Furthermore, in the study, although the researcher drew upon conversation analysis (CA) during the transcription process, a pure CA methodology was not applied while analyzing the results. Nevertheless, the analysis of interaction may require a pure conversation analysis methodology. Hence, the videos can be transcribed in a more detailed way applying the conventions suggested by

conversation analysts in further studies and nonverbal language can be included into the transcriptions to gain deeper insight into the use of interactional resources.

Lastly, in this study, only the role of proficiency and familiarity factors on interaction in paired tests were investigated. However, it is evident in the literature that there are other test-taker characteristics such as personality and gender which influence the test discourse. What is more, to the best knowledge of the researcher, there are no studies conducted on the role of personality and gender factors on interaction. Thus, the same study design can be used to investigate whether students' personality or gender play a role in EFL learners' use of interactional resources and the emergence of interactional patterns.

Conclusion

This descriptive study which was carried out with 100 EFL learners investigated whether interactant related factors (i.e., familiarity and proficiency) play a role in EFL learners' use of interactional resources and the emergence of interactional patterns in paired speaking tests. The findings yielded that pairing two different proficiency level students is disadvantageous for the high level test-takers in terms of topic management, task management and repair. On the contrary, while the low levels are at the advantageous side in terms of topic and task management in particular, they are at the disadvantageous side in terms of turn-taking. Moreover, the findings revealed that proficiency factor has a role on the interactional patterns generated by the test-takers. While high-high pairs create a collaborative pattern which is the most favorable one, low-low pairs usually create a parallel pattern. On the other hand, low-high pairs generate an asymmetric pattern due to the dominance of the high levels. Furthermore, the findings suggest that pairing two unfamiliar peer interlocutors seem more advantageous for the test-takers because when they are

familiar with each other they work more individually and there is the dominance of one test-taker in the exam. However, when they are unfamiliar with each other, they work more collaboratively. Therefore, while unfamiliar pairs generate a collaborative pattern, familiar pairs create an asymmetric pattern.

To the knowledge of the researcher, there are not many studies on the role of proficiency and familiarity factors on interaction in paired tests. Thus, these findings might shed light on the implementation of paired speaking tests. In conclusion, it is hoped that findings of this study and their pedagogical implications will contribute to the existing knowledge of the importance of the pairing system in paired tests and test administrators will gain insights into what factors they should take into consideration while pairing the test-takers in order not to destroy the reliability of the tests.

REFERENCES

Alderson, C., & Bachman, L.F. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.

Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F., & A. Palmer. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bachman, L.F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing*, *17*(1), 1-42.

Berry, V. (2004). Personality characteristics as a potential source of language test bias. In A. Huhta, K. Sahavaara, & S. Takala (Eds.), *Language testing: New openings* (pp. 115-124). Jyvaskyla, Finland: Institute for Educational Research.

Berry, V. (2007). *Personality differences and oral test performance. Language testing and evaluation*. Frankfurt am Main: Peter Lang.

Birjandi, P. (2011). From face-to-face to paired oral proficiency interviews: The nut is yet to be cracked. *English Language Teaching, 4*(2), 169-175.

Bonk, W.J., & Van Moere, A. (2004). *L2 group oral testing: The influence of shyness/outgoingness, match of interlocutors' proficiency level, and gender on individual scores.* Paper presented at the Language Testing Research Colloquium, Temecula, California.

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing, 26*(3), 341–66.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*.

New York: Longman.

Brown, A., & McNamara, T. (2004). “The devil is in the detail”: Researching gender issues in language assessment. *TESOL Quarterly, 38*(3), 524-538.

Burns, A., & Joyce, H. (1997). *Focus on speaking.* Sydney: National Center for English Language Teaching and Research.

Canale, M. & Swain, M. (1981). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.

Council of Europe. (2001). Common European framework of reference for languages: Learning, teaching, assessment. Cambridge: Cambridge University Press. Retrieved from http://www.coe.int/t/dg4/linguistic/source/framework_en.pdf

Csepes, I. (2002). Is testing speaking in pairs disadvantageous for students? A quantitative study of partner effects on oral test scores. *novELTy, 9* (1), 22-45.

Csépes, I. (2009). Measuring oral proficiency through paired-task performance. *Language testing and evaluation*. Frankfurt am Main: Peter Lang.

Cubillos, J. (2010). Computer-mediated oral proficiency assessments: Validity, reliability and washback. *International Journal of Technology, Knowledge and Society, 6*(6), 85–102.

Damon, W., & Phelps, E. (1989). Critical distinctions among three approaches to peer education. *International Journal of Educational Research, 13*(1), 9-19.

Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing, 26*(3), 367-396.

Dimitrova-Galaczi, E. (2008). Peer-peer interaction in a speaking test: The case of the First Certificate in English examination. *Language Assessment Quarterly, 5*(2), 89–119.

Dobao, A. F. (2012). Collaborative dialogue in learner–learner and learner–native speaker interaction. *Applied Linguistics, 33*(3), 229-256.

Ducasse, A. M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing, 26*(3), 423-443.

Ducasse, A. M. (2010). *Interaction in paired oral proficiency assessment in Spanish: Rater and candidate input into evidence based scale development and construct definition.* Frankfurt am Main: Peter Lang.

Drew, P. (1997). 'Open' class repair initiators in response to sequential sources of troubles in conversation. *Journal of Pragmatics, 28*(1), 69-101.

Egyud, G., & Glover, P. (2001).Oral testing in pairs: A secondary school perspective. *ELT Journal, 55*(1), 70-76.

Faerch, C., & Kasper, G. (1983). Plans and strategies in foreign language communication. In C. Faerch, & G. Kasper, (Eds.), *Strategies in interlanguage communication* (pp. 20-60). London: Longman.

Feldstein, S., & Welkowitz, J. (1987). A chronography of conversation: In defense of an objective approach. In A.W. Siegman, & S. Feldstein (Eds.), *Nonverbal behavior and communication* (2nd ed., pp. 435-499). Hillsdale, NJ: Erlbaum.

Fulcher, G. (2003). *Testing second language speaking*. Harlow: Pearson Longman.

Galaczi, E. D. (2004). *Peer-peer interaction in a paired speaking test: The case of the First Certificate in English* (Unpublished doctoral dissertation), Colombia University, New York. Retrieved from Teachers College.

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

Galaczi, E. D. (2014). Interactional competence across proficiency levels: How do learners manage interaction in paired speaking tests? *Applied Linguistics, 35*(5), 553-574.

Gan, Z., Davison, C., & Hamp-Lyons, L. (2008). Topic negotiation in peer group oral assessment situations: A conversation analytic approach. *Applied Linguistic, 30*(3), 315-334.

Gan, Z. (2010). 'Interaction in group oral assessment: A case study of higher- and lower-scoring students'. *Language Testing, 27*(4), 585–602.

Green, A. (2014). *Exploring language assessment and testing: Language in action.* Oxon: Routledge.

Greer, T., & Potter, H. (2012). Turn-taking practices in multi-party EFL oral proficiency tests. *Journal of Applied Linguistics, 5*(3), 297–320.

Hall, J. K. (1993). The role of oral practices in the accomplishment of our everyday lives: The sociocultural dimension of interaction with implications for the learning of another language. *Applied Linguistics, 14*(2), 145–166.

Hall, J. K. (1995). (Re)creating our worlds with words: A sociohistorical perspective of face-to face interaction. *Applied Linguistics*, *16*(2), 206–232.

Hall, J.K., & Pekarek Doehler, S. (2011). Introduction: Interactional competence and Development. In J.K. Hall, J. Hellermann, & S. Pekarek Doehler (Eds.), *L2 interactional competence and development* (pp.206-243). Clevedon: Multilingual Matters,

Hasselgren, A. (2000). The assessment of the English ability of young learners in Norwegian schools: an innovative approach. *Language Testing,17*. 261–277.

Ikeda, K. (1998). The paired learner interview: A preliminary investigation applying Vygotskyan insights. *Language Culture and Curriculum, 11*(1), 71-96.

Iwashita, N. (1996). The validity of the paired interview format in oral performance assessment. *Melbourne Papers in Language Testing, 5*(2), 51–65.

Iwasita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning : A Journal of Research in Language Studies, 51*(3), 401-436.

Jacoby, S., & Ochs, E. (1995). Co-construction: An introduction. *Research on Language and Social Interaction, 28*(3), 171-183.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. H. Lerner (Ed), *Conversation analysis: Studies from the first generation* (pp. 13-31). Amsterdam: John Benjamins.

Johnson, Marysia. 2000. Interaction in the oral proficiency interview: problems of validity. *Pragmatics, 10*(2), 215–31.

Johnson, M. (2001). *The art of nonconversation*. New Haven, CT: Yale University Press.

Kasper, G., & Ross, S. J. (2013). *Assessing second language pragmatics*. Basingstoke, Hampshire, UK: Palgrave Macmillan.

Kenyon, D. M., & Malabonga, V. (2001). Comparing examinee attitudes toward computer-assisted and other oral proficiency assessments. *Language, Learning and Technology, 5*(2), 60–83.

Kley, K. (2015). *Interactional competence in paired speaking tests: Role of paired task and test-taker speaking ability in co-constructed discourse* (doctoral dissertation). University of Iowa, Iowa, United States. Retrieved from: https://ir.uiowa.edu/cgi/viewcontent.cgi?article=5715&context=etd

Kormos, J. (1999). Simulating conversations in oral-proficiency assessment: A conversation analysis of role plays and nonscripted interviews in language exams. *Language Testing, 16,* 163–88.

Kramsch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal, 70*(4), 366-372.

Lazaraton, A. (2002). *A qualitative approach to the validation of oral language tests*. Cambridge, England: Cambridge University Press.

Lazaraton, A., & Davis, L. (2008). A microanalytic perspective on discourse, proficiency, and identity in paired oral assessment. *Language Assessment Quarterly, 5*(4), 313-335.

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press, Cambridge.

Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.

Markee, N. (2008). Toward a learning behavior tracking methodology for CA-for-SLA. *Applied Linguistics, 29*(3), 404-427.

May, L. (2009). Co-constructed interaction in a paired speaking test: The rater's perspective. *Language Testing, 26*(3), 397–421.

May, L. (2011). *Interaction in a paired speaking test*. Frankfurt am Main: PeterLang.

May, L. (2011). Interactional competence in a paired speaking test: Features salient to raters. *Language Assessment Quarterly, 8*(2),127–45.

Maynard, D. W. (1980). Placement of topic choices in conversation. *Semiotica, 30,* 263-290.

McNamara, T. F. (1996). *Measuring second language performance*. New York: Longman.

McNamara, T. F. (1997). Interaction' in second language performance assessment:

Whose performance? *Applied Linguistics, 18*(4), 446-466.

McNamara, T. F. (2001). Language assessment as social practice: Challenges for research. *Language Testing, 18*(4), 333-349.

McNamara, T., & Roever, C. (2006). *Language testing: The social dimension.* Oxford: Language Learning Monograph Series.

Messick, S. (1996). Validity and washback in language testing. *Language Testing 13*(3), 241-256.

Morrow, K. (1979) Communicative language testing: Revolution or evolution? In C.J. Brumfit, & K. Johnson (Eds.), *The communicative approach to language teaching* (pp.9–25). Oxford: Oxford University Press.

Nagano, R. (1997). Self-repair of Japanese speakers of English: A preliminary comparison with a study by W. J. M. Levelt. *Bulletin of Language Science and Humanities*, *11*, 65-90.

Nakatsuhara, F. (2011). Effects of test-taker characteristics and the number of participants in group oral tests. *Language Testing, 28*(4), 483–508.

Nakatsuhara, F. (2013). *The co-construction of conversation in group oral tests.* Frankfurt/Main, Germany: Peter Lang.

Norton, J. (2005). The paired format in the Cambridge speaking tests. *ELT Journal, 59*(4), 287-297.

Ockey, G. J. (2009). The effects of group members' personalities on a test-taker's L2 group oral discussion test scores. *Language Testing, 26*(2), 161-186.

Ockey, G. J., Koyama, D., & Setoguchi, E. (2013). Stakeholder input and test design: A case study on changing the interlocutor familiarity facet of the group oral discussion test. *Language Assessment Quarterly, 10*(3), 292-308.

Okada,Y. (2010). Role play in oral proficiency interviews: Interactive footing and

interactional competencies. *Journal of Pragmatics, 42*(6), 1647–1668.

Oller, J.W. (1979). *Language tests at school*. London: Longman.

O'Loughlin, K. (2002). The impact of gender in oral proficiency testing. *Language Testing, 19*(2), 169-192.

O'Sullivan, B. (2000). Exploring gender and oral proficiency interview performance. *System, 28*(3), 373-386.

O'Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing, 19*(3), 277-295.

O'Sullivan, B. (2002). *Modelling performance in tests of spoken language*. Frankfurt: Peter Lang.

Porter, D. (1991a). Affective factors in language testing. In J. C. Alderson, & B. North (Eds.), *Language testing in the 1990s. Modern English publications in association with the British Council* (pp.32-40). London: Macmillan.

Porter, D. (1991b). Affective factors in the assessment of oral interaction: Gender and Status. In S. Arnivan (Ed.), *Current developments in language testing* (pp.92-102). Singapore: SEAMEO Regional Language Centre. Anthology Series 25.

Sandlund, E., & Sundqvist, P. (2013). Diverging task orientations in L2 oral proficiency tests - a conversation analytic approach to participant understandings of pre-set discussion tasks. *Nordic Journal of Modern Language Methodology, 2*(1), 1–21.

Sandlund, E., Sundqvist, P., & Nyroos, L. (2016). Testing L2 talk: A review of empirical studies on second-language oral proficiency testing. *Language and Linguistics Compass, 10*(1), 14-29.

Savignon, S. J. (1972). *Communicative competence: An experiment in foreign language teaching.* Philadelphia, PA: Center for Curriculum Development.

Saville, N., & Hargreaves, P. (1999). Assessing speaking in the revised FCE. *ELT Journal, 53*(1), 42-51.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Linguistic Society of America, 53*(2), 361-382.

Schegloff, E. A. (2000). When 'others' initiate repair. *Applied Linguistics, 21*(2), 205-243.

Schegloff, E. A. (2007). *Sequence organization in interaction: Volume 1: A primer in conversation analysis.* Cambridge: Cambridge University Press.

Storch, N. (2001). *An investigation into the nature of pair work in an ESL classroom and its effect on grammatical development* (Unpublished Dissertation). University of Melbourne, Melbourne.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52* (1), 119-158.

Sun, D. (2014). From communicative competence to interactional competence: A new outlook to the teaching of spoken English. *Journal of Language Teaching and Research, 5*(5), 1062-1070.

Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing, 18*(3), 275–302.

Taylor, L. (2000). Investigating the paired speaking test format. *Cambridge ESOL Research Notes, 2*, 14–15.

Taylor, L., & Wigglesworth, G. (2009). Are two heads better than one? Pair work in L2 assessment contexts. *Language Testing*, *26*(3), 325-339.

Taylor, L. (2011). *Examining speaking: Research and practice in assessing second language speaking.* Cambridge: Cambridge University Press.

Van Lier, L. (1989). Reeling, writhing, drawling, sketching, and fainting in coils: Oral proficiency interviews as conversation. *TESOL Quarterly*, *23*(3), 489-508.

Van Lier, L. (1996). *Interaction in the language curriculum: Awareness, autonomy and authenticity.* London: Longman.

Van Moere, A. (2007). *Group oral tests: How does task affect candidate performance and test scores?* (Unpublished doctoral dissertation). Lancaster University, Lancaster, England.

Van Moere, A. (2014). Paired and group oral assessment. *Working Papers in TESOL and Applied Linguistics. 14*(2). Retrieved from: www.tc.columbia.edu/tesolalwebjournal

Walsh, S. (2011). Classroom interactional competence. In S. Walsh (Ed.), *Exploring classroom discourse: Language in action* (pp.158-181). New York: Routledge.

Wang, L. (2015). *Assessing interactional competence in second language paired speaking tasks* (doctoral dissertation). Northern Arizona University, Arizona, United States. Retrieved from: https://pqdtopen.proquest.com/doc/1708664834.html?FMT=ABS

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, *11* (2), 121-142.

Weir, C. J. (2005). *Language testing and validation*. Palgrave: Macmillan.

Widdowson, H.G. (1978). *Teaching language as communication*. Oxford: Oxford University Press.

Widdowson, H.G. (1979). *Explorations in applied linguistics*. Oxford: Oxford University Press.

Widdowson, H.G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.

Ying, B. (2009). The impact of familiarity on group oral proficiency testing. *CELEA, 32*(2), 114- 125.

Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition, 14*, 403-424.

Young, R. F. (1995). Conversational styles in language proficiency interviews. *Language Learning, 45*(1), 3–42.

Young, R., & He, A. W. (1998). *Talking and testing: Discourse approaches to the assessment of oral proficiency.* Amsterdam, Netherlands: John Benjamins Publishing Company.

Young, R. F. (2000, March). Interactional competence: Challenges for validity. In *Interdisciplinary interfaces with language Testing*. Symposium conducted at the Annual Meeting of the American Association for Applied Linguistics and the Language Testing Research Colloquium. Vancouver, Canada.

Young, R. F. (2008). *Language and interaction: An advanced resource book.* London, England: Routledge.

Young, R. (2011). Interactional Competence in Language Learning, Teaching and Testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 426-443). New York: Routledge.

APPENDIX A: Role-play Tasks

Task 1 Student A You work at a company. Today you are ill and you want to go to hospital. You call your company to talk to your boss, but your boss is not in his office so you talk to the secretary and leave your message. (You start the conversation.)	**Task 1** Student B You are the secretary. One of the workers is ill, so s/he wants to talk to the boss to tell him s/he will not be able to come to work because s/he wants to go to hospital. However, the boss is not in his office. You ask his/her personal information and get his/her message for the boss.

Task 2 Student A It's the end of the semester and you are planning a school picnic for the next week so you talk to your friend about the date, place, food and activities. However, next week can be rainy so you need to find an alternative place except school. (You start the conversation.)	**Task 2** Student B It's the end of the semester and your friend is planning a school picnic for the next week so you talk about the date, place, food and activities together to organize the picnic. However, next week can be rainy so you need to find an alternative place except school.

Task 3 Student A You are a customer at a restaurant. You order your meal and then you ask for the check, but when you get the check you realize that there is a problem with it so you talk to the waiter and ask for his/her help.	**Task 3** Student B You are a waiter at a restaurant. A customer comes and you take his/her order. Then s/he asks for the check. When s/he gets the check, s/he tells you there is a problem with it and asks for your help. (You start the conversation.)

Task 4 Student A You are ill so you go to a pharmacy. S/he asks you your symptoms and you tell him/her. You are allergic to aspirin so you cannot use it. You need to get another drug. After getting the drug, you pay for it. (You start the conversation.)	**Task 4** Student B You are a pharmacist. One patient comes and tells you s/he doesn't feel well so you ask his/her symptoms. S/he is allergic to aspirin so you need to give him/her another drug. After giving the drug, s/he needs to pay for it.

APPENDIX B: Duration of Videos

Video 1	04:17
Video 2	02:43
Video 3	07:43
Video 4	05:51
Video 5	02:13
Video 6	05:31
Video 7	04:27
Video 8	03:51
Video 9	03:35
Video 10	04:14
Video 11	07:08
Video 12	02:28
Video 13	05:11
Video 14	04:35
Video 15	06:21
Video 16	03:13
Video 17	02:51
Video 18	04:16
Video 19	05:03
Video 20	03:01
Video 21	04:50
Video 22	03:28
Video 23	04:01
Video 24	06:34
Video 25	03:55

Video 26	02:06
Video 27	04:15
Video 28	04:44
Video 29	04:56
Video 30	03:54
Video 31	05:25
Video 32	04:04
Video 33	05:43
Video 34	05:21
Video 35	03:42
Video 36	04:10
Video 37	03:42
Video 38	05:28
Video 39	03:04
Video 40	03:20
Video 41	02:58
Video 42	03:53
Video 43	03:57
Video 44	06:12
Video 45	02:56
Video 46	02:24
Video 47	04:23
Video 48	03:19
Video 49	03:52
Video 50	04:11

APPENDIX C: Rubric

ORAL PROFICIENCY TEST-Evaluation Sheet for the Assessor

A2	**Fluency & Pronunciation**	**Vocabulary Range**	**Grammatical Range & Accuracy**	**Spoken Production (Impromptu Speech)**	**Spoken Interaction (Role Play Task)**
4 ***Excellent to very good***	Can keep going comprehensively, even though pausing for grammatical and lexical planning and repair is very evident. Pronunciation is clear with occasional mispronunciations and a foreign accent is sometimes evident.	Shows good control of vocabulary when talking about familiar topics but errors still occur when dealing with unfamiliar topics. Attempts to paraphrase but with mixed success.	Good grammatical control. Occasional error and minor flaws may occur. Does not make mistakes which lead to misunderstanding.	Can reasonably fluently sustain a straightforward description of a familiar subject giving adequate details. Completes the task successfully.	Can understand clear, standard speech directed at him/her but s/he sometimes needs repetition and reformulation. S/he can usually understand enough to keep the conversation going.
3 ***Good to average***	Can communicate with relative ease but pauses, false starts are very noticeable. Pronunciation is generally clear enough to be understood with a noticeable foreign accent. Repetitions are necessary from time to time.	Has a sufficient vocabulary to talk about familiar everyday topics but errors are frequent when talking about less familiar topics. Rarely attempts paraphrase.	Can communicate with reasonable accuracy in familiar contexts but subordinate structures are rare Errors occur but it is usually clear what s/he is trying to say.	Can give a simple description of a familiar subject as a short series of simple phrases and sentences linked into a list. Can give most of the details and has moderate success in completing the task.	Can use some basic expressions in everyday/concrete situations. S/he is rarely able to understand enough to keep the conversation going.
2 ***Fair to poor***	Can manage very short isolated utterances with long pauses to search for expressions, correct pronunciation. Mispronunciations are frequent and cause some difficulty for the listener.	Has a basic repertoire of isolated words and phrases related to particular concrete situations. Has insufficient vocabulary control when talking about less familiar topics. Lack of vocabulary control sometimes leads to misunderstandings.	Produces basic sentence forms and some correct simple sentences but subordinate structures are rare Errors are frequent and may lead to misunderstandings.	Can talk about a familiar topic using simple isolated phrases. Limited success in completing the task. The description may include some irrelevant data.	Responds with short isolated phrases. Has problems in comprehension. Limited success in completing the task.
1 ***Very poor***	Pauses lengthily before most words. Little communication possible. Mispronunciations are numerous.	Communication is impaired from inadequate vocabulary	Cannot produce basic sentence forms Limited knowledge of sentence construction, does not communicate.	Can give very few details. The description is mostly irrelevant. Very little effort to complete the task.	Major problems in responding to her/his partner and in understanding the requirements of the task. Very little effort to complete the task.

- **If the speaker makes no attempt to respond OR response is IRRELEVANT to the topic, the speaker will get <u>1.</u>**

APPENDIX D: Transcription Conventions

[ ]	Square brackets mark the start and end of overlapping speech. They are aligned to mark the precise position of overlap as in the example below.
(0.4)	Numbers in round brackets measure pauses in seconds (in this case, 4 tenths of a second). If they are not part of a particular speaker's talk they should be on a new line. If in doubt use a new line.
(0.3)	A pause of 0.3 second
(1.0)	A pause of one second
(.)	A micro-pause, hearable but too short to measure
((stoccato))	Additional comments from the transcriber, e.g. about features of context or delivery
Yeh,	'Continuation' marker, speaker has not finished; marked by fall-rise or weak rising intonation, as when delivering a list.
Yeh.	Full stops mark falling, stopping intonation ('final contour'), irrespective of grammar, and not necessarily followed by a pause.
heh heh	Voiced laughter. Can have other symbols added, such as underlining, pitch movement, extra aspiration, etc.
sto(h)p i(h)t	Laughter within speech is signaled by h's in round brackets.
((unintelligible))	indicates a stretch of talk that is unintelligible to the analyst

APPENDIX E: Descriptive Statistics of Interactional Resources in Low-Low Pairs

APPENDIX E.1: Descriptive Statistics of Overlap in Low-Low Pairs

TurntakingOverlap

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By neither of them	15	100,0	100,0	100,0

APPENDIX E.2: Descriptive Statistics of Turn Speed in Low-Low Pairs

Turnspeed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	One test-taker is faster	3	20,0	20,0	20,0
	Both test-takers' speed is the same	12	80,0	80,0	100,0
	Total	15	100,0	100,0	

APPENDIX E.3: Descriptive Statistics of Turn Length in Low-Low Pairs

Turnlength

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	One test-taker has longer turns	8	53,3	53,3	53,3
	Both test-takers have the same turn length	7	46,7	46,7	100,0
	Total	15	100,0	100,0	

APPENDIX E.4: Descriptive Statistics of Turn Dominance in Low-Low Pairs

Turndominance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Dominance by one test-taker	7	46,7	46,7	46,7
	No dominance	8	53,3	53,3	100,0
	Total	15	100,0	100,0	

APPENDIX E.5: Descriptive Statistics of Topic Initiation in Low-Low Pairs

Topicinitiation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By the test-taker that the task requires	9	60,0	60,0	60,0
	By the other test-taker that the task doesn't require	4	26,7	26,7	86,7
	The examiner	2	13,3	13,3	100,0
	Total	15	100,0	100,0	

APPENDIX E.6: Descriptive Statistics of Topic Shift in Low-Low Pairs

Topicshift

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By one test-taker	4	26,7	26,7	26,7
	By the examiner	1	6,7	6,7	33,3
	No topic shift	10	66,7	66,7	100,0
	Total	15	100,0	100,0	

APPENDIX E.7: Descriptive Statistics of Topic Extension in Low-Low Pairs

Topicextension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By one test-taker	6	40,0	40,0	40,0
	By both test-takers	4	26,7	26,7	66,7
	By neither of them	5	33,3	33,3	100,0
	Total	15	100,0	100,0	

APPENDIX E.8: Descriptive Statistics of Topic Extension-2 in Low-Low Pairs

Topicextension2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Limited topic extension	10	66,7	66,7	66,7
	No topic extension	5	33,3	33,3	100,0
	Total	15	100,0	100,0	

APPENDIX E.9: Descriptive Statistics of Topic Closure in Low-Low Pairs

Topicclosure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By both test-takers	1	6,7	6,7	6,7
	By the examiner	14	93,3	93,3	100,0
	Total	15	100,0	100,0	

APPENDIX E.10: Descriptive Statistics of Task Management in Low-Low Pairs

Taskmanagement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By the examiner	2	13,3	13,3	13,3
	No task management	13	86,7	86,7	100,0
	Total	15	100,0	100,0	

APPENDIX E.11: Descriptive Statistics of Repair in Low-Low Pairs

Repair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes when it is required	5	33,3	33,3	33,3
	No even if it is required	4	26,7	26,7	60,0
	No because not required	6	40,0	40,0	100,0
	Total	15	100,0	100,0	

APPENDIX F: Descriptive Statistics of Interactional Resources in High-High Pairs

APPENDIX F.1: Descriptive Statistics of Overlap in High-High Pairs

TurntakingOverlap

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By one test-taker	5	33,3	33,3	33,3
	By both test-takers	2	13,3	13,3	46,7
	By neither of them	8	53,3	53,3	100,0
	Total	15	100,0	100,0	

APPENDIX F.2: Descriptive Statistics of Turn Speed in High-High Pairs

Turnspeed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	One test-taker is faster	4	26,7	26,7	26,7
	Both test-takers' speed is the same	11	73,3	73,3	100,0
	Total	15	100,0	100,0	

APPENDIX F.3: Descriptive Statistics of Turn Length in High-High Pairs

Turnlength

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	One test-taker has longer turns	9	60,0	60,0	60,0
	Both test-takers have the same turn length	6	40,0	40,0	100,0
	Total	15	100,0	100,0	

APPENDIX F.4: Descriptive Statistics of Turn Dominance in High-High Pairs

Turndominance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Dominance by one test-taker	7	46,7	46,7	46,7
	No dominance	8	53,3	53,3	100,0
	Total	15	100,0	100,0	

APPENDIX F.5: Descriptive Statistics of Topic Initiation in High-High Pairs

Topicinitiation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By the test-taker that the task requires	15	100,0	100,0	100,0

APPENDIX F.6: Descriptive Statistics of Topic Shift in High-High Pairs

Topicshift

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By one test-taker	12	80,0	80,0	80,0
	By the examiner	3	20,0	20,0	100,0
	Total	15	100,0	100,0	

APPENDIX F.7: Descriptive Statistics of Topic Extension in High-High Pairs

Topicextension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By one test-taker	2	13,3	13,3	13,3
	By both test-takers	13	86,7	86,7	100,0
	Total	15	100,0	100,0	

APPENDIX F.8: Descriptive Statistics of Topic Extension-2 in High-High Pairs

Topicextension2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Good topic extension	10	66,7	66,7	66,7
	Limited topic extension	5	33,3	33,3	100,0
	Total	15	100,0	100,0	

APPENDIX F.9: Descriptive Statistics of Topic Closure in High-High Pairs

Topicclosure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By one test-taker	6	40,0	40,0	40,0
	By both test-takers	6	40,0	40,0	80,0
	By the examiner	3	20,0	20,0	100,0
	Total	15	100,0	100,0	

APPENDIX F.10: Descriptive Statistics of Task Management in High-High Pairs

Taskmanagement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No task management	15	100,0	100,0	100,0

APPENDIX F.11: Descriptive Statistics of Repair in High-High Pairs

Repair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes when it is required	6	40,0	40,0	40,0
	No because not required	9	60,0	60,0	100,0
	Total	15	100,0	100,0	

APPENDIX G: Descriptive Statistics of Interactional Resources in Low-High Pairs

APPENDIX G.1: Descriptive Statistics of Overlap in Low-High Pairs

TurntakingOverlap

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By low level test-taker	1	5,0	5,0	5,0
	By high level test-taker	1	5,0	5,0	10,0
	No overlap	18	90,0	90,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.2: Descriptive Statistics of Turn Speed in Low-High Pairs

Turnspeed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low level test-taker is faster	1	5,0	5,0	5,0
	High level test-taker is faster	14	70,0	70,0	75,0
	Both test-takers' speed is the same	5	25,0	25,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.3: Descriptive Statistics of Turn Length in Low-High Pairs

Turnlength

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low level test-taker has longer turns	1	5,0	5,0	5,0
	High level test-taker has longer turns	18	90,0	90,0	95,0
	Both test-takers have the same turn length	1	5,0	5,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.4: Descriptive Statistics of Turn Dominance in Low-High Pairs

Turndominance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Dominance of high level test-taker	17	85,0	85,0	85,0
	No dominance	3	15,0	15,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.5: Descriptive Statistics of Topic Initiation in Low-High Pairs

Topicinitiation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By the low level TT as the task requires	4	20,0	20,0	20,0
	By the high level TT as the task requires	13	65,0	65,0	85,0
	By the high level TT although the task requires the low level TT	1	5,0	5,0	90,0
	By the examiner	2	10,0	10,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.6: Descriptive Statistics of Topic Extension in Low-High Pairs

Topicextension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By high level TT	15	75,0	75,0	75,0
	By both of them	5	25,0	25,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.7: Descriptive Statistics of Topic Extension-2 in Low-High Pairs

Topicextension2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Good topic extension	5	25,0	25,0	25,0
	Limited topic extension	15	75,0	75,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.8: Descriptive Statistics of Topic Shift in Low-High Pairs

Topicshift

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By high level TT	8	40,0	40,0	40,0
	By low level TT	5	25,0	25,0	65,0
	By both of them	3	15,0	15,0	80,0
	No topic shift	4	20,0	20,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.9: Descriptive Statistics of Topic Closure in Low-High Pairs

Topicclosure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By low level TT	3	15,0	15,0	15,0
	By high level TT	7	35,0	35,0	50,0
	By both of them	1	5,0	5,0	55,0
	By the examiner	9	45,0	45,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.10: Descriptive Statistics of Task Management in Low-High Pairs

Taskmanage

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By high level TT	13	65,0	65,0	65,0
	By the examiner	1	5,0	5,0	70,0
	No task management	6	30,0	30,0	100,0
	Total	20	100,0	100,0	

APPENDIX G.11: Descriptive Statistics of Repair in Low-High Pairs

Repair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes by high level TT when it is required	13	65,0	65,0	65,0
	No by low level TT even if it is required	2	10,0	10,0	75,0
	No because not required	5	25,0	25,0	100,0
	Total	20	100,0	100,0	

APPENDIX H: Descriptive Statistics of Interactional Patterns in Low-Low Pairs

InteractionalPatterns

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	asymmetric	1	6,7	6,7	6,7
	parallel	10	66,7	66,7	73,3
	blend(parallel+asymmetric)	4	26,7	26,7	100,0
	Total	15	100,0	100,0	

APPENDIX I: Descriptive Statistics of Interactional Patterns in High-High Pairs

InteractionalPatterns

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	collaborative	7	46,7	46,7	46,7
	asymmetric	1	6,7	6,7	53,3
	parallel	1	6,7	6,7	60,0
	blend(collaborative+asymmetric)	6	40,0	40,0	100,0
	Total	15	100,0	100,0	

APPENDIX J: Descriptive Statistics of Interactional Patterns in Low-High Pairs

InteractionalPatterns

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	collaborative	2	10,0	10,0	10,0
	asymmetric	11	55,0	55,0	65,0
	blend(collaborative+asymmetric)	4	20,0	20,0	85,0
	blend(parallel+asymmetric)	3	15,0	15,0	100,0
	Total	20	100,0	100,0	

APPENDIX K: Descriptive Statistics of Interactional Resources in Familiar Pairs

APPENDIX K.1: Descriptive Statistics of Overlap in Familiar Pairs

TurntakingOverlap

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	There is overlap	4	44,4	44,4	44,4
	No overlap	5	55,6	55,6	100,0
	Total	9	100,0	100,0	

APPENDIX K.2: Descriptive Statistics of Turn Speed in Familiar Pairs

TurnSpeed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	4	44,4	44,4	44,4
	Equal	5	55,6	55,6	100,0
	Total	9	100,0	100,0	

APPENDIX K.3: Descriptive Statistics of Turn Length in Familiar Pairs

TurnLength

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	7	77,8	77,8	77,8
	Equal	2	22,2	22,2	100,0
	Total	9	100,0	100,0	

APPENDIX K.4: Descriptive Statistics of Turn Dominance in Familiar Pairs

TurnDominance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	8	88,9	88,9	88,9
	Equal	1	11,1	11,1	100,0
	Total	9	100,0	100,0	

APPENDIX K.5: Descriptive Statistics of Topic Initiation in Familiar Pairs

TopicInitiation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By the test-taker that the task requires	9	100,0	100,0	100,0

APPENDIX K.6: Descriptive Statistics of Topic Extension in Familiar Pairs

TopicExtension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	4	44,4	44,4	44,4
	Equal	5	55,6	55,6	100,0
	Total	9	100,0	100,0	

APPENDIX K.7: Descriptive Statistics of Topic Shift in Familiar Pairs

TopicShift

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	9	100,0	100,0	100,0

APPENDIX K.8: Descriptive Statistics of Topic Closure in Familiar Pairs

TopicClosure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	6	66,7	66,7	66,7
	Equal	1	11,1	11,1	77,8
	No topic closure	2	22,2	22,2	100,0
	Total	9	100,0	100,0	

APPENDIX K.9: Descriptive Statistics of Task Management in Familiar Pairs

TaskManagement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	2	22,2	22,2	22,2
	No task management	7	77,8	77,8	100,0
	Total	9	100,0	100,0	

APPENDIX K.10: Descriptive Statistics of Repair in Familiar Pairs

Repair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes when it's required	4	44,4	44,4	44,4
	No because not required	5	55,6	55,6	100,0
	Total	9	100,0	100,0	

APPENDIX L: Descriptive Statistics of Interactional Resources in Unfamiliar Pairs

APPENDIX L.1: Descriptive Statistics of Overlap in Unfamiliar Pairs

TurntakingOverlap

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	There is overlap	2	22,2	22,2	22,2
	No overlap	7	77,8	77,8	100,0
	Total	9	100,0	100,0	

APPENDIX L.2: Descriptive Statistics of Turn Speed in Unfamiliar Pairs

TurnSpeed

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	4	44,4	44,4	44,4
	Equal	5	55,6	55,6	100,0
	Total	9	100,0	100,0	

APPENDIX L.3: Descriptive Statistics of Turn Length in Unfamiliar Pairs

TurnLength

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	5	55,6	55,6	55,6
	Equal	4	44,4	44,4	100,0
	Total	9	100,0	100,0	

APPENDIX L.4: Descriptive Statistics of Turn Dominance in Unfamiliar Pairs

TurnDominance

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	5	55,6	55,6	55,6
	Equal	4	44,4	44,4	100,0
	Total	9	100,0	100,0	

APPENDIX L.5: Descriptive Statistics of Topic Initiation in Unfamiliar Pairs

TopicInitiation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	By the test-taker that the task requires	9	100,0	100,0	100,0

APPENDIX L.6: Descriptive Statistics of Topic Extension in Unfamiliar Pairs

TopicExtension

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	2	22,2	22,2	22,2
	Equal	7	77,8	77,8	100,0
	Total	9	100,0	100,0	

APPENDIX L.7: Descriptive Statistics of Topic Shift in Unfamiliar Pairs

TopicShift

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	5	55,6	55,6	55,6
	Equal	1	11,1	11,1	66,7
	No topic shift	3	33,3	33,3	100,0
	Total	9	100,0	100,0	

APPENDIX L.8: Descriptive Statistics of Topic Closure in Unfamiliar Pairs

TopicClosure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	3	33,3	33,3	33,3
	Equal	4	44,4	44,4	77,8
	No topic closure	2	22,2	22,2	100,0
	Total	9	100,0	100,0	

APPENDIX L.9: Descriptive Statistics of Task Management in Unfamiliar Pairs

TaskManagement

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unequal	1	11,1	11,1	11,1
	No task management	8	88,9	88,9	100,0
	Total	9	100,0	100,0	

APPENDIX L.10: Descriptive Statistics of Repair in Unfamiliar Pairs

Repair

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes when it's required	4	44,4	44,4	44,4
	No because not required	5	55,6	55,6	100,0
	Total	9	100,0	100,0	

APPENDIX M: Descriptive Statistics of Interactional Patterns in Familiar Pairs

InteractionalPatterns

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	collaborative	2	22,2	22,2	22,2
	asymmetric	3	33,3	33,3	55,6
	blend(collaborative+asymmetric)	3	33,3	33,3	88,9
	blend(parallel+asymmetric)	1	11,1	11,1	100,0
	Total	9	100,0	100,0	

APPENDIX N: Descriptive Statistics of Interactional Patterns in Unfamiliar Pairs

InteractionalPatterns

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	collaborative	4	44,4	44,4	44,4
	asymmetric	2	22,2	22,2	66,7
	blend(collaborative+asymmetric)	3	33,3	33,3	100,0
	Total	9	100,0	100,0	

APPENDIX O: Low Level Test-takers' Scores

Low Level Test-taker	Oral Proficiency Exam Score	Proficiency Exam Score
1	6	38
2	7	39
3	5	45
4	5	36
5	7	44
6	5	38
7	9	43
8	8	40
9	8	46
10	7	41
11	9	44
12	5	42
13	10	42
14	9,5	35
15	6	30
16	10	36
17	8	38
18	8	42
19	7,5	41
20	10	39
21	8	44
22	7	44
23	5	40
24	5	38
25	8	42
26	6	39
27	7,5	41
28	9	42
29	7	42
30	7	40
31	5	40
32	5	38
33	9,5	34
34	7,5	36
35	5,5	29
36	8	33
37	8	35
38	7	41
39	6	38
40	6	39
41	9	42
42	8	37

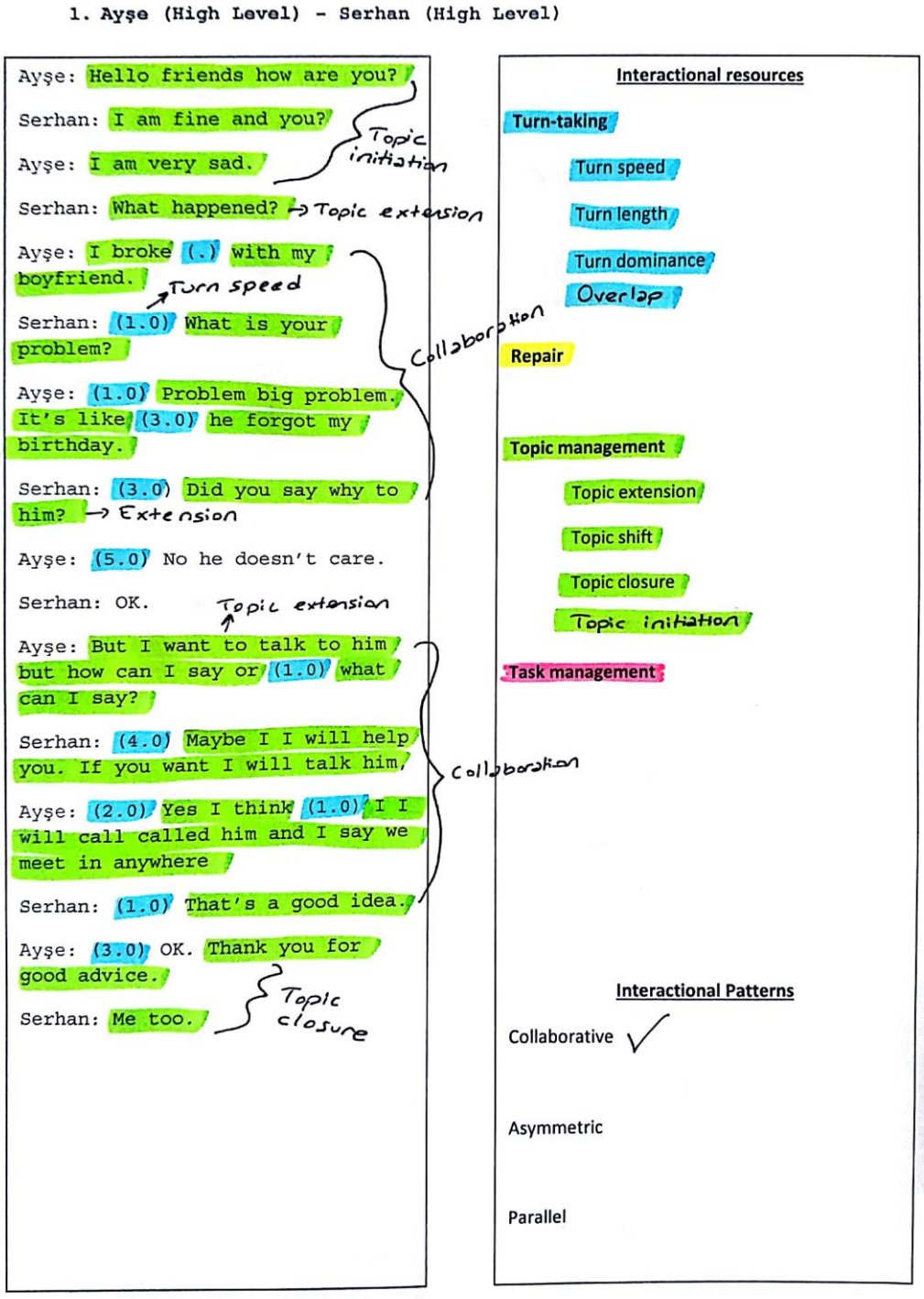
43	9	40
44	8	34
45	10	40
46	8	29
47	7,5	38
48	6	30
49	8	35
50	9	

APPENDIX P: High Level Test-takers' Scores

High Level Test-taker	Oral Proficiency Exam Score	Proficiency Exam Score
1	20	65
2	18	62
3	20	61
4	19,5	58
5	19	63
6	19	67
7	18,5	56
8	16,5	55
9	16,5	62
10	17	62
11	17	63
12	17	65
13	18	70
14	16	65
15	17	58
16	16	63
17	17	59
18	15	61
19	18	57
20	16,5	72
21	18	66
22	17	65
23	20	65
24	17	60
25	20	63
26	15	58
27	19,5	64
28	19,5	65
29	18	61
30	16,5	66
31	19,5	73
32	19,5	60
33	19	57
34	20	72
35	19	69
36	20	67
37	19	66
38	18	57
39	19	59
40	17	62
41	20	68
42	20	65

43	20	73
44	17	59
45	17	70
46	18,5	64
47	18	67
48	18	61
49	20	67
50	20	69

APPENDIX R: Color-coding



APPENDIX S: A Sample of the Collaborative Pattern

Gizem (High Level)- Faruk (High level)

Gizem: Welcome

Faruk: Welcome

Gizem: How can I help you?

Faruk: (2.0) I don't feel better because I have a flu, (1.0)

Gizem: What is your (1.0) symptom or problem?

Faruk: (3.0) I have a headache and cold, (2.0)

Gizem: OK (1.0) have you (1.0) any allergic to drugs?

Faruk: Yes madam. I allergic to drugs aspirin.

Gizem: OK (2.0) I I give you penicillin but you should to see doctor

Faruk: OK (3.0) how much do the drugs today? (1.0) [Or day] days.

Gizem: [Eee (Well)] a day (.) two.

Faruk: (2.0) OK if I don't feel better (1.0) I go to (1.0) or (.) I should (.) see a doctor or I go to (1.0) pharmacy?

Gizem: I think you can (.) go to doctor.

Faruk: OK (3.0) how much is it?

Gizem: It's 5.99 dollar.

Faruk: Heh heh OK OK please madam. ((He gives the money))Let's see you later. Thank you.

Gizem: Thank you.

APPENDIX T: A Sample of the Asymmetric Pattern

Burak (Low Level)- Taner (High Level)

Taner: Hello welcome
Burak: Welcome
Taner: Would you like to drink something or would you like to drink some meal?(3.0)
Burak: Water(.) şey (well)
Taner: You are in the restaurant now ((He warns his partner))
Burak: Yes restaurant Şey (4.0) Water cola.
Taner: (.) OK two or one water? (1.0) One, how many cola do you want?(6.0)
Burak: One please.
Taner: One cola OK I will bring for you(6.0)OK and then your your (2.0) your bill (2.0) 25 dollars you know (1.0) your bill 25 dollars.
Burak: 25 dollars
Taner: Yeah. (8.0)
((Burak doesn't understand and wants to leave the exam))
Taner: I will give the box here is the box you need to put pay money.
Burak: Şey money
Taner: Money yeah but there is a problem it is very expensive.
((The task so indicates that))
Burak: Şey what is your problem?
Taner: Your problem. The bill is very expensive for you you know. (1.0) A lot of money.
Burak: A lot of money
((They cannot go on since Burak doesn't understand what he needs to say))

APPENDIX U: A Sample of the Parallel Pattern

D**amla (Low Levlel)- Hakan (Low Level**

Damla: Do you have a childhood memories?
Hakan: ((He nods his head as he does not understand the question))
Damla: Do you have a childhood memories? ((She repeats the question))
Hakan: Yes, I do.
Damla: Wha..?
(30.0)
((No more attempt))

APPENDIX V: A Sample of the Blend (Collaborative+Asymmetric) Pattern

Uras (High Level)- Yağız (High Level)

Uras: So you know whole year we were here some school exam and,
Yağız: Yeah.
Uras: School will finish soon,
Yağız: Hopefully.
Uras: Yeah. (1.0) According to that we should do some things activities with friends,
Yağız: Of course of course we choose (unintelligible).
Uras: We should spend time together with [the others],
Yağız: [Yes].
Uras: If it is possible for them.
Yağız: It's not in a good day you know and maybe we can do a picnic.
Uras: General generally (unintelligible).
Yağız: I think we are planning together like this maybe we can do picnic on park you know,
Uras: Yeah that's a [great idea].
Yağız: [Green places] you know,
Uras: It's cool it's really cool.
Yağız: Yes. What about you? What do you think about it?
Uras: Yeah (1.0) I don't know I am I you know I am from Adana so we should make a mangal (grill) kebap (kebab),
Yağız: Yes of course maybe we can buy chicken or meat you know,
Uras: Oh I am gonna (going to) ask you something you know,
Yağız: Yeah of course.
Uras: I heard something about the Muslim people,
Yağız: Heh heh
Uras: Yo(h)u kno(h)w this is big problem to eat pork for you? Ha? Do you believe that?
Yağız: Eee (Well),
Uras: So if I bring it can we eat together? I want you to eat
Yağız: It depends people to people but we are eating together something all of the class so maybe we can just buy a chicken,
Uras: [Yeah].
Yağız: [And] we can find a place you know,
Uras: [Alright].
Yağız: [Forest] or something.
Uras: So we should call the friends and next week is alright?
Yağız: Of course alright maybe 10 am is OK but,
Uras: Are they drinking normal or without (.)[ex],
Yağız: [Yeah].
Uras: excluded alcohol?
Yağız: I don't know but we are doing do do we can do something together of course it depends people to people [maybe],
Uras: [Oh]!
Yağız: Some people don't want this.
Uras: There is a picnic saloon you know place,
Yağız: Of course.
Uras: I can meet there for the picnic if there is a big it might be rainy [so it might rain],
Yağız: [It isn't going to be bad] it isn't going to be bad maybe it's it's happen I don't know. And maybe we can play football or basketball something if we can (.)[you know],
Uras: [Aaa (well)],
Yağız: Court or something.
Uras: Yeah yeah we can.
Examiner: That's enough. Thank you.

APPENDIX Y: A Sample of the Blend (Parallel+Asymmetric) Pattern

Olcay (Low level)- Hale (Low Level)

((Before starting the conversation Hale states that she doesn't understand the task and the examiner explains it. Although she is supposed to start the conversation by asking her partner some questions about his childhood, she could not, so Olcay starts the conversation.))
Olcay: When I was child (1.0) I don't like go to school. (5.0) My(6.0) şey ((well)). (2.0) My child friend name Arda. Who is friend?
(2.0)
Hale: Hıı (Yeah) where did you go?
(13.0)
Examiner: Answer his question please. He asked you a question. Who was your best friend?
(4.0)
Olcay: Who is best friend name?
Hale: (3.0) My best friend is İpek.
Olcay: How old are you?
Hale: She, (9.0)
Examiner: What did you do with her when you were a child?
Hale: (4.0) I play I played game, (5.0) I went to İzmir.
(4.0)
Examiner: Yes, would you like to add something about your childhood? ((The examiner poses the question to Olcay))
Olcay: (2.0) Repeat?
Examiner: Your best friend, the games you played when you were a child?
Olcay: (1.0) My best friend Arda. Arda he is (2.0) 20 years old. He (6.0) Dokuz Eylül University (2.0) Electrical Electronic department engineering. (13.0) I like him (1.0) because (3.0) I (1.0) a long time with (3.0) him, (5.0) were (where) do you live my best friend?
Hale: (3.0) My best friend live in Ankara.
Olcay: (4.0) What type of music do you like? (2.0) Best friend.
Hale: (3.0) She favorite songs pop music.
Examiner: Yes that's enough. Your exam is over.