Interaction Between Speakers in Oral Language Proficiency Test

Anna Suur-Kujala
English department
University of Helsinki
Master's Thesis
18.2.2007
Table of Contents

1. Introduction .....................................................................................................1

2. Theoretical framework.....................................................................................4
   2.1. The development of the notion of language proficiency .......................4
       2.1.1 Communicative competence .................................................................4
       2.1.2 Towards the interactional model ......................................................... 5
       2.1.3. The CEF on interaction in oral proficiency assessment .................... 7
   2.2. Testing and assessing oral language proficiency ...................................... 9
       2.2.1. Test validity and reliability ................................................................. 9
       2.2.2. Oral language proficiency testing .....................................................12
       2.2.3. Previous studies on oral proficiency testing .................................... 13
   2.3. Oral proficiency test formats ...................................................................16
       2.3.1. Individual interview or discussion ....................................................17
       2.3.2. Paired interviews ............................................................................17
       2.3.3. Group task ........................................................................................19
   2.4. Turn-taking .............................................................................................19
       2.4.1. Conversation analysis and turn-taking ..............................................19
       2.4.2. Turn-taking in this study ..................................................................23

3. Research question and hypothesis ..................................................................25

4. Materials and methods ...................................................................................26
   4.1. Grouping of the speakers .......................................................................28
1. Introduction

These days, language proficiency is regarded to be much more than just the knowledge of grammatical rules or lexicon. When reaching the 21st century, the concept of language proficiency has gradually moved towards communicative and interactive competence. This shift of priorities - or, rather the emerging of yet another priority – has raised new kinds of demands. As the spoken language and oral communication have become one of the centrepoints of language teaching, the testing and assessment of these skills have also received more attention.

However, testing and assessing oral proficiency is not a simple matter. The questions of general reliability and validity of any test are present in this situation as well. Furthermore, the test results is affected by factors that do not occur in a written test: the raters’ values and opinions, the test situation and the other people (such as other interlocutors) involved in it as well as the person’s personal and social abilities. Testing and assessing speaking is a challenging field of study. Luoma (2004: 190) points out that we need to make sure we understand the different aspects involved in an interactive situation such as pair or grouped interview. She adds that the understanding of how a person’s performance influences the performance of others should be a central question in future studies.

There are several factors which affect the assessment of oral proficiency. In a pair discussion, the keys to one’s success do not entirely lie in one’s hands as the communicative situation includes two people. Therefore, the impact of one interlocutor on the result of the other needs to be carefully considered when
assessing oral language skills. The more we gain knowledge about the interlocutors’ impact on the outcome, the more valid the oral language proficiency assessing can become.

My study is conducted as a part of the HY-Talk project which is funded by the University of Helsinki. It is carried out by the Department of Applied Sciences and the language departments. The group includes students, teachers and researches. HY-Talk studies the assessment of spoken language and oral language proficiency. Its aim is to improve the reliability and validity of assessment procedures of spoken language and as well as the CEFR (Common European Framework of Reference for Languages) based proficiency scales modified for the Finnish school system.

The HY-Talk test tasks are relatively structured and only give a limited amount of freedom to the speakers. Is there nevertheless differentiation in the distribution of turn sizes when they are faced with different partners? The amount of speech is an important factor in oral language proficiency tests as it is hard to assess the test-takers if they are not talkative enough. The HY-Talk group of professionals who evaluated the language proficiency of the students who took part in the study, found it difficult to assess them as some of them were talking so little. It can be argued that if the raters do not receive enough data, the evaluation of the student’s language skills is not likely to be valid.

This master's thesis concentrates on studying interaction in a pair test situation, and on how the dynamics and interaction changes in relation to the interlocutor.
The analysis of the interaction is done by using turn-taking as the theoretical framework. The size of the turns is measured in each conversation. The change in the size of turns is compared, when the interlocutor is paired with people with different language skills. The level of each of the interlocutor’s language skills is evaluated according to oral proficiency scales included in the Finnish national core curriculum for language education.
2. Theoretical framework

The theoretical framework of this thesis consists of three main sections. In the first one, I will talk about the change in the notion of language proficiency; how the communication and interaction are today’s key concepts in language education. The second part will discuss the spoken language testing and assessment and the problems that arise from it. The third part will introduce the term turn-taking as it is central concept in this thesis.

2.1. The development of the notion of language proficiency

In this chapter I discuss the notion of language proficiency by looking at some of the most influential notions that has been developed on the topic. Canale and Swain (1980) discussed communicative competence and their construction of the concept has become the basis of the whole communicative view of language and influenced language teaching and learning a great deal during the past few decades. I will also introduce two other models which reflect the gradual change to interactional models of language proficiency.

2.1.1 Communicative competence

Hymes (1971) coined the concept communicative competence, and from the 1970s on, the concept of language ability and proficiency has gradually changed from knowing the core facts of language (grammar and lexicon) towards communicativity and interaction. In today’s world, foreign languages are the key
to international and intercultural communication. In this chapter, I define the concept of language ability by introducing a few influential and central approaches to the topic.

Canale and Swain (1980) were pioneers in SLA in regarding communicative skills as part of language competence. They divide the components of communicative competence into three groups:

1) Grammatical competence: the knowledge of correct grammar rules in a language
2) Sociolinguistic competence: to know how to use socially correct and acceptable language
3) Strategic competence: the ability to use communication strategies in order to utter the intended meaning.

2.1.2 Towards the interactional model

Bachman (1990) and Bachman and Palmer (1996) have also developed Canale and Swain’s definition further. They put emphasis not only on the language knowledge but also on the language use, that is the communicative aspect of language. The term interaction occurs in the discussion, and even though they still regard language testing as merely testing of an individual’s abilities, they acknowledge the complexity of different interactional processes involved in it.

Bachman and Palmer (1996: 61-62) state that “-- in general, language use can be defined as the creation or interpretation of intended meanings in discourse by an
individual, or as the dynamic and interactive negotiation of intended meanings between two or more individuals in a particular situation”. They state that the language use situation is a combination of several complex and intertwining interactions that occur among the individual characteristics of language users as well as in the actual language use situation. Thus they present an interactional framework for approaching the concept of language ability.

Bachman and Palmer (1996: 62-63) suggest that the interacting components which together form language use are “language ability (language knowledge and strategic competence, or metacognitive strategies), topical knowledge and affective schemata” and personal characteristics such as age, sex and native language. These components interact with the characteristics of language use or test task and setting. This means that the different circumstances and demands present in the language use situation contribute to an individual’s language ability and they should be taken into account when assessing tests which measure language use.

Like Bachman and Palmer, Celce-Murcia et al. (1995: 17-18) also revised Canale and Swain’s views of communicative competence and took interaction as one of the key elements of language ability. They recognised the significance of all the participants of the interactive situation. Celce-Murcia et al. include five areas in the concept of communicative competence. They define communicative competence as consisting of linguistic competence, strategic competence, sociocultural competence, actional competence and discourse competence. In this model, Canale and Swain’s sociolinguistic competence is divided into two groups:
sociocultural competence and actional competence. Rather than seeing interaction as a joint effort, Celce-Murcia et al. speak of the different aspects of interaction mostly from an individual’s viewpoint and characteristics (i.e. from turn-taking to the understanding of the cultural differences) but their notion of sociocultural competence also recognises the impact of the interlocutor. Communication is not something that two participants of a discussion, for instance, create from their own standpoints but rather a joint construct where they both affect each other and the outcome of the communicative situation.

Celce-Murcia et. al (1995:3) have described some of these variables affecting sociocultural competence. They called them social contextual factors. These factors take into account the participants in the interactive situation. The interlocutor’s age, gender, profession and a number of other variables affect the way they communicate and how other people communicate with them. Thus, Celce-Murcia et al. (1995) do not see the language ability as merely an individual performance but rather a complex process where all the participants affect each other’s performances.

2.1.3. The CEF on interaction in oral proficiency assessment

The Common European Framework of Reference for Languages (CEF) (Council of Europe, 2001: 125) was created as to provide a common ground and framework for European language education and assessment. The CEF also recognises the significance of interaction in communication. It states that language competence is not merely just a question of knowing what language
form is used to convey a specific language function. Language use is about interaction. The CEF (2001: 14) states that “in interaction at least two individuals participate in an oral and/or written exchange in which production and reception alternate and may in fact overlap in oral communication. - - Learning to interact thus involves more than learning to receive and to produce utterances. High importance is generally attributed to interaction in language use and learning in view of its central role in communication.”

In the discussion of oral language proficiency evaluation, The CEF (2001: 193) provides 12 qualitative categories which are relevant to the evaluation. The categories deal with different aspects of spoken language, and six of these categories either directly or partly call for paying attention to the interactive language use. These categories are:

- Turn-taking strategies
- Co-operating strategies
- Asking for clarification
- Thematic development
- Flexibility
- Sociolinguistic competence

However, even though the CEF recognises the interactive nature of a oral language usage, it does not emphasise the effect of the interlocutor in the situation or talk about the assessment of oral language proficiency as a joint effort but rather the individual’s effect on the communicative situation. However, these
2.2. Testing and assessing oral language proficiency

In this section, I will discuss the validity and reliability of a language test. The assessment of a test should be based on reliable measurements and the variables affecting the result should be recognised. Furthermore, this section introduces the different forms of oral language testing and discusses the problems that can occur in spoken language assessment. I will also take a look at some previous studies conducted on interlocutor’s impact on the oral language test result.

2.2.1. Test validity and reliability

Creating a valid and reliable test is a result of delicate consideration. Bachman and Palmer (1996: 9) discuss the fundamental principles of language testing. For the test to serve its purpose meaningfully, it is worthwhile paying attention to these different aspects.

The first one of these principles has to do with the correspondence between language test performance and actual language use. Bachman and Palmer (1996: 10-11) point out that if the aim is to measure a person’s language skills, the test
needs a conceptual framework which ensures that the test performance is a close representation of non-test language use. The characteristics of the circumstances (what kind of language use situation is in question and who are the ones using the language or taking the test) should be carefully considered.

Another fundamental principle for language testing that Bachman and Palmer (1996: 17-39) present is test usefulness. They list six qualities which help to determine the usefulness of a given test: reliability, construct validity, authenticity, interactiveness, impact and practicality. I will briefly clarify each of the qualities.

Reliability at its shortest can be defined as “consistency of measurement”. Bachman and Palmer (1996: 19) point out that ultimately, the reliability of the test ensures that the test score remains consistent regardless of the circumstances in which it is carried out. The scores given by two different raters in the same test should not vary from each other, nor should the results of two different tests measuring the same ability.

Construct validity, according to Bachman and Palmer (1996: 21) refers to the correctness of assumptions that we make about the test results. To what extent are the interpretations accurate? When constructing a test, one should ensure that it really measures the ability in question. Furthermore, Weir (1993: 45) emphasises that the beginning of all test-related factors is always the test validity. What is a relevant and convenient way to gather and assess the data? These matters need to be taken into consideration in order to establish reliable results. However, Bachman and Palmer (1996: 21) point out that test validity is a continuous
process. and it is present in all stages of test construction and use, but it is unlikely that it will ever be absolute.

The third aspect that Bachman and Palmer (1996: 23) raise is that of authenticity. They argue that the level of correspondence between a test task and a target language use task is the key question in authenticity. This quality helps to evaluate the degree to which the test scores can be generalised beyond the domain of test taking situation. In other words, do the test scores provide accurate information of target language use in non-test situations. Furthermore, Bachman and Palmer claim that the authenticity of the test has a positive effect on the test-takers as they perceive it as relevant and reasonable way to test language skills.

The term interactiveness, on the other hand, consists of the test taker’s personal abilities that influence the result of a test task. According to Bachman and Palmer (1996: 25), the most important individual characteristics for language testing are “language ability (language knowledge and strategic competence, or metacognitive strategies), topical knowledge and affective schemata”. The concept of language proficiency is discussed in the previous section of this thesis.

The last two qualities that Bachman and Palmer (1996: 30-36) list as crucial elements of a language test, are impact and practicality. The former refers to the impact that the test has on society and educational systems, and on the individual people who are part of these systems. Tests do not occur in a vacuum but rather reflect the values, mentality and goals of a society and educational system. Practicality, in turn, serves a different purpose from the other qualities. Whereas
the other qualities deal with the implications and effects of the test scores, practicality has to do with the ways the test is used, and if it is used at all. The resources available have to be taken into consideration before the implementation of a test is possible.

2.2.2. Oral language proficiency testing

Huhta (1993: 79, 84) points out that as communicative language skills became an essential part of language testing, new problems arose with it. During the behaviouristic era, language tests were largely multiple choice oriented as that test form guaranteed a certain reliability in the assessment. Later, when the communicative aspects of language began to gain ground in language teaching, the tests underwent a transformation as well.

Huhta (1993:85) continues that communicative language tests such as oral language proficiency test do not provide as clear-cut answers about the candidate’s language skills as traditional tests. A language is used in a social situation and thus there are several non-linguistic factors which can affect the outcome of a test. Therefore the assessment of oral skills is more than mechanically counting the points in a multiple choice test. It is a qualitative evaluation which is a more challenging form of assessment.

In his short review of history of oral proficiency testing, Huhta (1993: 143) states that even though oral language testing was initially regarded as unreliable and ambiguous, it has gradually improved its reputation. There has been a shift from discussing the reliability of assessment to the validity of the test itself. Huhta
defines validity as an evaluation of how accurately and reliably the test measures the features that it is intended to measure.

An important factor in the validity of the test is the assessment criteria. Huhta (1993: 149-155) points out that a careful consideration of what different aspects the examiners need to pay attention to is the key to reliable and valid testing. In oral language assessment, there are two different paths which give the guidelines to this consideration. Holistic assessment pays attention to the whole speech act and does not distinguish between different aspects of language (e.g. pronunciation, grammar). By contrast, an analytic approach divides the speech into smaller units which together form the evaluation of the speech. Huhta states that while holistic assessment is widely used, especially due to its more straightforward nature (and to the fact that it is also faster to carry out and thus less expensive), it may lack the elaborateness that analytic assessment can, at least to some extent, provide.

Of relevance to this thesis is Huhta’s notion (1993: 172) that many oral proficiency tests (e.g. Cambridge Assessment of Spoken English) have listed communication strategies and interaction as one of the assessment criteria.

2.2.3. Previous studies on oral proficiency testing

Luoma (2004: 12, 16-17) discusses the criteria that are often considered when assessing the oral proficiency of an L2 learner. She notes that it is common to judge one’s speaking skills according to one’s lexicon and the correct use of
grammar, i.e. how vividly and correctly the learner can express his/her thoughts. These two factors receive the main attention when assessing speech.

The effect of the test-taker’s characteristics on the language test result has been studied e.g. by Kunnan (1995) and Purpura (1998). These studies evaluated the effect of different factors on the test results. These factors included e.g. education level, motivation, exposure to the target language. The studies showed that they play a significant role in the variation in test-taker performances.

The interlocutor’s effect on the result has been studied somewhat but McNamara (1996: 85-86) criticises the oral language test performance studies in general for not including the interaction between the interlocutors in them. He (1996: 87) adds that “language ability is to be assessed under performance conditions, that is, as part of and act of communication”. Weir (2005: 107-108) comments on the same problem. He states that the interaction in oral proficiency test situation is in fact co-constructed as there are at least two people involved, but the grades go to each individual separately.

O’Sullivan (2002: 277) is one of the people who have studied the interlocutor’s effect on the outcome of a test. He notes that the oral test result is influenced by several factors that are related to the test-taker, the partner and the task itself. He himself has studied learner acquaintanceship (2002) and come to the conclusion that the familiarity of the interlocutor has a positive effect on the oral proficiency test result. Other variables studied from the interlocutor’s perspective are age
Weir (1993: 35) has also analysed spoken interaction tests. According to him, the outcome does not only depend on the language skills of the examinee but on a number of other factors as well. Among these other variables Weir mentions the time that the examinee has to prepare to the test. The preparation time given should be in relation to how long turns the speaker is expected to take. A lecture or speech naturally requires more time than spontaneous reacting to an everyday situation.

Another factor affecting the outcome of an oral proficiency test has to do with the reciprocity conditions. Weir (1993: 35-37) states that the speaker-listener relationship and the interactional dimension of the communicative situation should be taken into account when assessing oral skills. He, like O’Sullivan, argues that familiarity between the participants in a spoken test situation may well make the situation more comfortable and even easier. Furthermore, it may be easier to talk to a friend or a person of the same age than to a (possibly unknown) authority like the examiner. This relaxation factor influences the language sample that can be elicited. A relaxed examinee will speak more freely and thus more information can be gathered about his/her oral proficiency.

A study close to mine was conducted in Australia by Iwashita in 1997. She carried out a study where paired interview was used in order to evaluate the test result in relation to the interlocutor’s oral language proficiency. The twenty
students, who took part in the study, were all native English speakers with a few bilinguals whose second language was Chinese. They were all studying Japanese. In her study, Iwashita found out that the test-takers tended to receive higher grades in the test when their interlocutor’s language skills were at the higher end of the proficiency scale than when they were in the lower end. She also discovered that in general the test-takers spoke more when their interlocutor was of higher proficiency. However, she (1997, 64) notes that this was not always the case as there were individual differences between the talkativeness of the students.

In her study, Iwashita (1997) used tasks with very little regulation: the speakers did not have as highly controlled content in the tasks as there is in the HY-Talk tasks.

In this chapter, I have discussed both language testing in general and oral language testing. The questions of validity and reliability are present in all testing and there are no clear-cut answers to these problems. However, more and more qualitative study is being conducted on e.g. the interaction in a spoken language proficiency test, and the rising demand is bound to result in new and helpful findings in the field.

2.3. Oral proficiency test formats

This chapter introduces common test formats which are in used in oral language proficiency testing. It briefly discusses the pros and cons of each one and
addresses some problems facing this area of testing in general, e.g. the lack of resources.

2.3.1. Individual interview or discussion

Luoma (2004: 35) writes that there are three methods for oral proficiency testing. According to her, the most common one is the individual interview format where the examinees are all assessed one at a time in separate test situations. This format became dominant in the 1950s when oral assessment gained more attention.

There are both strengths and weaknesses to the individual interview. Luoma (2004: 35) notes that the method allows a great deal of flexibility when it comes to the questions and topics of the test. They can be adapted to suit the examinee’s level of proficiency and this, in turn, gives the interviewer more power in the situation.

The power of the interviewer, however, can also be considered a weakness of the method. (Luoma 2004: 35) The interviewer takes the initiative in all phases of the discussion, and the examinee only replies to the questions. The one-on-one test situation can be more open if it is in the form of a conversation or discussion.

2.3.2. Paired interviews

Paired interviews differ from individual interviews in the interaction with the examiner. In a paired interview the two examinees interact mostly with each other and the role of the examiner is that of a rater. There are some advantages to this
testing format. Shohamy (1986, cited in Swain 2001: 277) notes that the paired interview offers the possibility to use different kinds of talk formats instead of the traditional interview. Thus it can provide the examiner with more information on the examinees oral proficiency. Furthermore, Hildson (1991, cited in Swain 2001: 277) argues that it may be beneficial to the learners to use the same models in testing that have been used in the classroom situation. Since pair discussions are often practiced during the language lessons, the relationship with teaching and testing can be emphasised in this test format. Thirdly, as Berry (2000, cited in Swain 2001: 277) states, the paired interview saves resources as it does not take as much of the examinees time as the individual interview.

In pair test situations, however, the dynamics is different as the responsibility to take initiative and responsibility of the interaction is more balanced between the interlocutors. Luoma (2004: 35) argues that the pair interaction situations do provide the raters with information about the examinees’ language skills. However, Weir (1993: 55) names some potential problems that might occur in a student-student spoken test situation. Weir argues that problems may develop in reciprocity conditions if one of the examinees takes control over the interaction. This may diminish the other person’s opportunity to fully prove his/her language competence. Furthermore, if the two examinees’ proficiency in the language differ significantly, it may have an effect on the performance and thus affect the assessment. Luoma (2004: 37) also mentions the familiarity factor (discussed earlier in section 2.2.3.) as something that should be taken into account with pair tasks. She emphasises that the result of a paired test has to be analysed with care
since there are several factors, besides the actual language skills, which may affect the outcome.

2.3.3. Group task

The third form of oral proficiency testing is a group task which consists of a group discussion. Luoma (2004: 39) notes that in a classroom situation this format can be practical but as a test format it requires delicate attention as the examiner has to ensure that all the participants are granted enough time to perform their language skills.

2.4. Turn-taking

In this section, I will introduce conversation analysis (CA) as a field of study and define the concept of turn-taking as part of it. I will also clarify how turn-taking is used in my study. In this thesis, I do not concentrate on the viewpoints of CA but rather use turn-taking as a tool for analysing my data.

2.4.1. Conversation analysis and turn-taking

Markee (2000: 64) writes that the field of conversation analysis (CA) is concerned with revealing the social competences which can be found in social interaction. Psathas (1995: 2) states that “conversation analysis studies the order/organization/orderliness of social action, particularly those social actions that are located in everyday interaction - -“. Therefore he suggests that the term
Conversation refers to too narrow an area and the field of study could well be called interaction analysis.

In my paper interaction is defined as the change that occurs in the turn sizes when the interlocutor changes. Thus, in this study, I will use CA’s concept of turn size to analyse the talkativeness of the speakers. Conversation analysis uses turn-taking for analysing the structure of talk-in-interaction. According to Markee (2000: 82), communication analysis sees turn-taking as a clarification of how the speakers together organise and cooperate in a communicative situation. Crookes (1990: 185) defines a turn as “one or more streams of speech bounded by speech of another, usually an interlocutor”. This definition, however, does not provide a straightforward model when analysing speech with multiple interlocutors and overlapping of speech. Sacks et al. (1974: 696-702) note that turn-taking is “a prominent type of social organization, one whose instances are implicated in a wide range of other activities”. It is, they add, essential to the organization of any conversation.

Different speech exchange systems (dialogue, lecture, a mass demonstration etc.) consist of different turn-taking conventions. Markee (2000: 87) points out that the more structured the speech exchange system is (e.g. a debate) the more structured the speech is. Speakers speak when they have the right to do so and they do it for a certain period of time.

The HY-Talk exercises used in this study fall under the category of organised speech exchange systems. They are structured as they are written in a form of a
dialogue and thus make the distribution of turns clear-cut. However, the tasks do not strictly control the size of the turns but as the contents of each line is given; there is not much room for free speech. The only exception is the last part of the test where the speakers are only given hints of what they should talk about and the order and more detailed contents of turns etc. are left for the speakers to decide.

Sacks et al. (1974: 700-701) have listed 14 characteristics of turn-taking. These characteristics can be identified in equal power speech exchange systems. Before discussing the characteristics, it is relevant to define the notion of an equal power speech exchange. Markee (2000: 68) distinguishes equal and unequal power speech exchange systems by the kind of rights the participants have for talking. In a teacher-student talk, for example, the teacher has the right to both speak and distribute turns to the students. Such a speech exchange system is unequal. By contrast, an ordinary conversation between friends is an equal power speech exchange situation. In this study, I categorise a paired oral proficiency test as equal power speech exchange situation. Firstly, the two participants are in an equal power situation: neither of them has more privileges to speak than the other. Secondly, even though the turns are distributed beforehand (as the participants are given roles and they must therefore follow the order in which their role characters speak), it does not strictly limit the size of a turn and in the freer pair discussion part the participants do not have to follow any specific order of turns. Therefore, the interlocutors can, within a certain framework, decide the size (and in some tasks the order) of their turns. Furthermore, even though the content of the participants’ speech is regulated, all the instructions are given in another language
than the target one. Thus, the participants are free to choose whichever way they want to utter the intended meaning.

According to Sacks et al. (1974: 700-701), the 14 characteristics of turn-taking that can be identified in a equal power speech exchange system list are:

1. Speaker-change recurs, or at least occurs.
2. Overwhelmingly, one party talks at a time.
3. Occurrences of more than one speaker at a time are common but brief.
4. Transitions (from one turn to a next) with no gap and no overlap are common. Together with transitions characterized by slight gap or slight overlap, they make up a vast majority of transitions.
5. Turn order is not fixed, but varies.
6. Turn size is not fixed but varies.
7. Length of conversation is not specified in advance.
8. What parties say is not specified in advance.
9. Relative distribution of turns is not specified in advance.
10. Number of parties can vary.
11. Talk can be continuous or discontinuous.
12. Turn-allocation techniques are obviously used. A current speaker may select a next speaker or parties may self-select in starting to talk.
13. Repair mechanisms exist for dealing with turn-taking errors and violations, e.g. if two parties find themselves talking at the same time, one of them will stop prematurely, thus repairing the trouble.
2.4.2. Turn-taking in this study

In this study, I pay attention to the turn sizes by which I mean counting all words (including articles and prepositions) uttered by a speaker in English. I do not, however, include backchannelling or hesitations in the count because they do not provide the kind of information about one’s language proficiency that the raters could use in evaluating a speaker’s English language skills. The sizes of turns are a relevant variable in measuring the interaction between the interlocutors in my study as they are straight forward to measure and provide exact and clear results in a speaker’s talkativeness. In my study the size of turns equals the talkativeness of each interlocutor and is thus a valid variable when measuring the change that occurs in the interaction when the interlocutor changes.

The rules of turn-taking are often studied from a cultural perspective. According to different studies (e.g. Lehtonen & Sajavaara 1982, Berry 1997), the turn-taking styles (e.g. talkativeness) can be characterised as a cultural phenomenon. People from different cultures act differently in communicative situations. In my study, however, I am more interested in individual differences and interaction in a paired oral proficiency testing situation.

Traditionally, the key question in turn-taking studies has had to with how individuals act and survive in communicative situations, i.e. how they recognise a point when to talk or when to listen. In my study, however, I use the size of turns as measurement of the interaction between the interlocutors. How do one’s language skills affect the amount his/hers interlocutor talks? The concept of turn-
taking provides a framework in which it is possible to measure the amount of speech that each person utters.

In this chapter I have introduced theoretical viewpoints which are relevant for my study: the notion of language proficiency, the problems and studies conducted on language testing and the tools for analysing my data provide the background information needed for understanding the meaning and intention of this study. In the next chapters I will establish the research questions and hypothesis and try to answer them according to the results I found when analysing the data.
3. Research question and hypothesis

The research questions are:

How does the interaction between speakers change in relation to the interlocutor’s oral language proficiency in spoken test tasks?

* What is the effect of a more proficient interlocutor?

* What is the effect of a less proficient one?

My hypothesis is that the students will speak more when faced with a more proficient interlocutor. I base this statement on Iwashita’s study (1997) and my own perceptions of language interaction. In Iwashita’s study, the students spoke more when they were faced with a more competent speaker although she notes that there were some variation (in some individual cases the more proficient interlocutor did not result in increased turn sizes).
4. Materials and methods

This study is qualitative by nature and could be called a case study since I analyse each participant in detail. The first phase of the study included 8 students aged 14 and 15. They are all studying in ordinary comprehensive school in Finland. I chose the students randomly but they are in the same class. I paired them randomly and they were asked to perform an oral task together. All the discussions were videotaped and then transcribed (Appendices 2-5).

After the performances, the eight students’ oral language proficiency was assessed according to the Finnish national core curriculum proficiency scales on their overall oral English language proficiency on a scale from 1-10, ten being the highest grade. The performances were assessed by a group of trained raters such as English teachers and university English students. Altogether 6 raters took part in the assessment. The interlocutors were assessed holistically that is their language skills were evaluated as a whole without dividing them into smaller units such as grammar or fluency, for example.

For the second phase of my study, I chose four of the eight students and regrouped them with a new partner. The pairs performed the same tasks again and they were again videotaped. After transcribing the second discussions, I calculated the words uttered by each student and compared the changes that occurred when they were faced with two differently proficient interlocutors.
The tasks used in the study are HY-Talk material. As this study was conducted as pilot study for the HY-Talk project, it was most logical to use its material for it. The project is using the performance of very specific tasks as its data, so therefore my I decided to collect my data as using the same tasks (Appendix 1). The tasks cover vocabulary and communicative situations which, according to the Finnish national core curriculum, the 14-year-old students are expected to have learnt in school. They are designed for the language proficiency level A1.3. The task consists of three sections of which the first two are in the form of a dialogue. The third part of the task is a freer dialogue-type discussion where the interlocutors are asked to negotiate a plan for a day out in Helsinki.

The eight 8th graders were randomly selected to take part in my study. Their level of oral English language proficiency was not evaluated beforehand and neither did I consult the teacher in this matter. The idea with this random selection was that of any given oral language test situation: the interlocutors are not being matched according to their language skills but rather they have to cope with whomever they are facing in the test situation.

The students were paired randomly and given 15 minutes to prepare for the tasks by themselves (no cooperation or taking notes was allowed). The time period between the two taping sessions was 3 months (summer holiday), and in the second taping session, the students said that they had hardly any recollection of the tasks they performed before summer. Thus, if the performance of the task was better in the second taping, it was not the result of remembering the task from the previous taping session. Also, I asked the students if they had been abroad (to an
English-speaking country) in order to find out if their oral language proficiency could have improved significantly but none of them had travelled there.

4.1. Grouping of the speakers

According to the evaluations, there was not a great deal of variation between the students' English language proficiency. It was possible, however, to distinguish between some of the weakest and strongest performances. But most of the students did fall under the same skill level, and the differences were only minor. The grades given to the students are presented in Table 1.

Table 1.

<table>
<thead>
<tr>
<th>Grouping</th>
<th>Student No</th>
<th>The level of oral proficiency (1=lowest 10=highest)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair A</td>
<td>Student 1</td>
<td>5,3</td>
</tr>
<tr>
<td>Pair A</td>
<td>Student 2</td>
<td>5,4</td>
</tr>
<tr>
<td>Pair B</td>
<td>Student 3</td>
<td>5,1</td>
</tr>
<tr>
<td>Pair B</td>
<td>Student 4</td>
<td>6,2</td>
</tr>
<tr>
<td>Pair C</td>
<td>Student 5</td>
<td>3,9</td>
</tr>
<tr>
<td>Pair C</td>
<td>Student 6</td>
<td>4,9</td>
</tr>
<tr>
<td>Pair D</td>
<td>Student 7</td>
<td>5,2</td>
</tr>
<tr>
<td>Pair D</td>
<td>Student 8</td>
<td>5,6</td>
</tr>
</tbody>
</table>

As the differences between most of the students were fairly insignificant, I decided to take pairs B and C and mix them because these two pairs included there most variation in the grades of these four students. Pairs B and C included
the strongest English speaker of the whole group as well as the two weakest. Furthermore the fourth speaker is also at the weaker end of the group but as the differences are minimal, I consider him a medium-level speaker (when compared with the rest of the group). As the two weakest students had been coincidentally paired together, I now paired the second weakest one with the best one and the medium-level speaker with the weakest one.

The two pairs are shown in Table 2.

**Table 2.**

<table>
<thead>
<tr>
<th>New grouping</th>
<th>Student No</th>
<th>The level of oral proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair D</td>
<td>Student 4</td>
<td>6,2</td>
</tr>
<tr>
<td>Pair D</td>
<td>Student 6</td>
<td>4,9</td>
</tr>
<tr>
<td>Pair E</td>
<td>Student 3</td>
<td>5,1</td>
</tr>
<tr>
<td>Pair E</td>
<td>Student 5</td>
<td>3,9</td>
</tr>
</tbody>
</table>
5. Results

In this section, I analyse the results of my study. I discuss the students’ task performances in detail. I have named one heading after each student but as the discussions always included two students, it is naturally impossible to separate them completely for analysis. Thus, all the students are discussed after each heading even though the main emphasis is always on one of them.

5.1. Comparison of the performances

Tables 3-6 I show the numbers of words each student uttered in the two discussions. In the brackets are the grades that each student received when their English language proficiency was evaluated.

Table 3. The number of words Student 3 uttered in the two discussions.

<table>
<thead>
<tr>
<th>Student 3 (5,1)</th>
<th>Discussion with Student 4 (6,2)</th>
<th>Discussion with Student 5 (3,9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words uttered</td>
<td>119</td>
<td>164</td>
</tr>
</tbody>
</table>

Table 4. The number of words Student 4 uttered in the two discussions.

<table>
<thead>
<tr>
<th>Student 4 (6,2)</th>
<th>Discussion with Student 3 (5,1)</th>
<th>Discussion with Student 6 (4,9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words uttered</td>
<td>283</td>
<td>220</td>
</tr>
</tbody>
</table>

Table 5. The number of words Student 5 uttered in the two discussions.

<table>
<thead>
<tr>
<th>Student 5 (3,9)</th>
<th>Discussion with Student 3 (5,1)</th>
<th>Discussion with Student 6 (4,9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words uttered</td>
<td>95</td>
<td>110</td>
</tr>
</tbody>
</table>
Table 6. The number of words Student 6 uttered in the two discussions.

<table>
<thead>
<tr>
<th>Student 6 (4,9)</th>
<th>Discussion with Student 4 (6,2)</th>
<th>Discussion with Student 5 (3,9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words uttered</td>
<td>153</td>
<td>138</td>
</tr>
</tbody>
</table>

As the Tables above show, there are differences in the word distribution between the participants. A medium-level speaker, student 3, was first paired with student 4 who had the strongest English skills. During the task he uttered 119 English words. However, when he was paired with the student with the weakest language skills, his number of words rose to 164. So, the medium-level speaker student 5 spoke notably more with the weakest English speaker than with the best one.

Student 4, with the strongest spoken English proficiency was fairly talkative with a medium-level speaker, student 3, with 283 words. When she was paired with a slightly weaker student, the number of words uttered decreased to 220 words.

Student 5 received the lowest grade of the group. She was first paired with student 6 who similarly was at the weaker end of English speakers in that group. She was later matched with the a person with slightly stronger proficiency – student 3 – but it did not have a significant effect on the number of words she uttered. The discussion with student 6 resulted in 110 words and the discussion with student 3 in 95 words.

Student 6 was another weak student in the group although closer to the middle-level speakers than the weakest one. She was first paired with the weakest speaker
– student 5 – and uttered then 138 words. When she was paired with the best English speaker, student 4, her result rose to 153 words.

5.2. Analysis of the discussions

Having listed the results in the previous section, I will now analyse them in more detail in the following paragraphs. I will analyse each student separately and compare the two discussions they have had. As there are two people in each discussion, it is impossible to analyse only one interlocutor at a time. Thus, all students are discussed in all sections but the main focus is always on the student whose number is on the heading.

5.3.1. Student 5

The person whose results showed the least variation was Student 5. She received the lowest grade and she remained the most silent during both conversations of which one was with lower proficiency speaker and one with the middle-level speaker. She actually spoke more with the less proficient speaker then with the middle-level speaker (the change was from 110 words to 95 words).

The raters who evaluated the students had difficulties in rating this particular student. They felt that they were not able to form a clear picture of her language skills based on the tasks she performed. This is a problem that should be solved by lengthening the tasks. If the students themselves do not take responsibility for speaking enough, the exercises should be designed in a way that they will provide enough material. Even highly structured tasks as in HY-Talk, make it possible to
speak very little. The tasks are written in the form of dialogue but it does not automatically solve the problem with silent speakers.

The case of Student 5 immediately raises other questions about her personal characteristics. It does not seem far-fetched to guess that her poor language skills contribute to her silence but that is not the only factor. The personality plays an important role. A shy or insecure student does not perform in the same way as an outgoing and talkative one. Although the speaker's shyness may well be the result of poor language skills and then the lower grade is the correct evaluation.

There are, however, other factors that may can contributing to the silence or what may seems shyness. People who suffer from test anxiety or otherwise feel uncomfortable in the test situation might not be able to show their strongest skills in that particular situation. The videotaping also might have made the student more nervous than she would have been in a more natural situation.

5.3.2. Student 6

Student 6 had quite the opposite results than Student 5. She was also graded as a fairly weak English student with the average grade of 4,9. While talking with another weak student (student 5), the size of her turns remained low. She did speak more than Student 5 which was the expected result regarding that she was evaluated as a more proficient speaker than her interlocutor. However, it seems that both Student 5 and Student 6 adopted the same kind of model of speaking: none of them was willing to speak more than was absolutely necessary, and I
assume that the quietness of Student 5 was partly transferred to Student 6 because in her second discussion, Student 6 was more talkative.

When Student 6 was regrouped with the best speaker in the group, Student 4, the outcome was different. Student 6’s size of turns increased from 135 to 153 words when she performed the tasks together with Student 4. It seems that Student 6 adapted the style of Student 4: she talked more and used longer utterances. She did not take the leading role in either of the two discussions but rather reacted according to her interlocutor. If the other person seemed more talkative, Student 6 spoke more as well. But when the interlocutor did not say more than was absolutely necessary, she also became more silent.

The data that was gathered from Student 6 suggests that having a more proficient English speaker as an interlocutor had a positive outcome on her performance. It seems that she felt more encouraged when her interlocutor spoke more as well.

5.3.3. Student 4

Student 4 received the highest score of 6.2 and she was unquestionably the most fluent and proficient English speaker of the group. Her turn sizes were on both occasions over 200 words, which is more than those of any other speaker. She was first paired with the middle-level student 3 whose average grade was 5.1 and then with a slightly weaker student 6 whose grade was 4.9. In the former discussion she uttered 283 words and in the latter one 220. The interlocutor’s language skills were fairly close to each other but there was a decrease in the turn sizes. It seems that Student 6, even though she was fairly silent when speaking
with the middle-level speaker student 3, did now take space from Student 4. Student 3, on the other hand, did not take an active role in the discussion which resulted in Student 4 speaking more than twice as much as him.

It is interesting that two students whose language skills are fairly close to each other (Students 3 and 6) performed differently when speaking with a highly proficient speaker. One of them stepped aside and did not contribute to the discussion more than was necessary whereas the other became more talkative. This observation suggests that there might be other factors than language proficiency that contribute to the result of the performance. Underlying social relations and personal features, for example, may well play a major role in the interaction between the two interlocutors. O’Sullivan (2002) states that familiarity between the interlocutors has an impact on the outcome of a pair discussion. However, he studied the familiar interlocutors versus strangers. In this case, all the students know each other as they are in the same class. Nevertheless, it is likely that there are underlying social relations and dynamics between interlocutors and the matter should be studied further.

5.3.4. Student 3

Student 3 was first paired with Student 4 and then with Student 5. Unlike I had assumed beforehand, speaking with a more proficient interlocutor did not have a positive effect on the middle-level Student 3. On the contrary, Student 3 appeared quite silent when faced with the more fluent English speaker. However, when he was paired with a silent and less proficient Student 5, he seemed to take charge of the communicative situation as his number of words uttered rose from 119 to 164.
It seems that Student 3 found it easier or more natural to talk in the discussion between him and a weaker student than between him and high-graded student. Student 3 compensated the quietness of the weaker student by being more active himself but in the other discussion he became more quiet as the high-graded student was talkative and controlled the interaction.
6. Discussion

In contrast to Iwashita’s (1997) study, the HY-Talk tasks used here were quite structured and controlled. They were in the form of a dialogue which helps with the more equal distribution of turns and turn sizes. In the free speech tasks that Iwashita used, turn-taking is not regulated and therefore it can be assumed that the more talkative speakers can easily control the discussion. In the HY-Talk exercises, on the other hand, one could expect a fairly equal distribution of turn sizes as the interlocutors had specific turns when to talk. However, surprisingly, there were major differences between the speakers’ turn sizes even when they used the HY-Talk exercises. Even though the contents of the speakers’ utterances were regulated, some of the students created more abundant and lively utterances, and expressed themselves more freely and this created variation in the turn sizes.

This small-scale study shows that interaction between the two speakers is not solely dominated by their or their interlocutor’s oral proficiency. On the contrary, the results show that the language skills – in their traditional sense (e.g. knowledge of grammar and vocabulary etc.) – may be a minor factor in the forming of interaction. There are several other factors that contribute to the result in such test situation. Firstly, the personalities of each interlocutor play a tremendous role in this form of testing. How shy, active, open, introvert, talkative or self-confident is the person who is put in this situation? And possibly even more importantly, how do the interlocutors bring out these features in each other? A shy student can feel intimidated when facing a talkative student, and an
insecure student can be encouraged by an even more insecure speaker and thus become more talkative.

Secondly, the relationship between the two speakers should be taken into account. This is a difficult and time-consuming task but its effect should not be underestimated. O’Sullivan (2002) states that acquaintanceship makes the communicative situation easier and thus results in better outcomes in similar settings to that used in my study. However, if all the interlocutors know each other, the details of their relationships become important. The group of students that I analysed was chosen from one class and this is often the case in oral language proficiency testing anyway. The students all knew each other since they have been in the same class for at least two years. However, their relationships were not examined closer. It is quite possible that some of them were friends, but on the other hand one of them could be a bully and one of them his/her “victim”. The personal relationships between the interlocutors may have a profound influence on the outcome of a test or a task where they have to interact with each other.

Furthermore, phenomena such as test anxiety may well result in poor test performance. The speaker might be nervous and not used to being evaluated for speaking a foreign language (or for any oral performance for that matter). The rater may make the test situation more uncomfortable just by being there, and videotaping may increase the nervousness of the speaker. I feel that test anxiety is something the raters could try to decrease in the oral language test situation. A small warm-up task or general small-talk before the actual performance might
help the speakers to relax and feel less nervous. On the other hand, how well one performs in any test and is able to control the nervousness is an asset that some people have and some do not. I would like to argue that it is hard to eliminate the stress factor altogether and create test situations that would be relaxed, casual and yet effective that there will always be poor results that were caused by a person being nervous.

The purpose of this study was to measure the change in the interaction when the interlocutor changes. I did not measure the differences in test results in relation to the interlocutor but rather the change in talkativeness. Iwashita (1997: 62) points out that talking more does not always result in a better test performance but I argue that studying it is still relevant for several reasons: Firstly, the fact that there was change in the student’s activeness reveals the complexity which underlies the interactive situation and it should be studied in more detail. Secondly, even though the change in their talkativeness might not directly affect their speaking performance as such, it does however affect the way the raters view the students and thus may have an impact on their test score. A more talkative student can be viewed as more proficient one even though it might not be the right evaluation. Even though being talkative is by no means any more or any less a virtue than being a silent, it may result in a better outcome in the test situation.

The results of this small study did not unambiguously confirm Iwashita’s (1997) results. The outcome of her study was that the students spoke more when faced with a more proficient interlocutor, and I had a reason to expect a similar result. There were, however, individual variation in Iwashita’s study. As my data was
small and only meant for qualitative analysis, it could be argued that my results fall under Iwashita’s category of “individual differences”, that is, that my results could be seen as an exception to the rule. However, I would like to argue that even though the data is small, it passes on a valid point to the discussion of oral language proficiency assessment. It shows, more than anything, that the interaction between then interlocutors (and thus the outcome of a oral language test, too) is a mixture of several intertwining factors, and that, even more importantly, language proficiency might not be among the most important ones. Further study and a better understanding of other underlying factors (such as the relationship between the interlocutors) seems crucial if we want to gain understanding of the interaction process as a whole. This understanding, in turn, is essential in our attempts to steer oral language proficiency assessment to the direction that e.g. Bachman and Palmer (1996) and Luoma (2004) discuss: to make as valid, reliable and objective as possible.
7. Conclusion

The small-scale study described in this thesis resulted in further questions. First and foremost, I feel that a study of the students’ relationships would provide useful information on why the results are what they are. Interviewing the students after their performance could enlighten the underlying social issues that can affect the result.

Furthermore, it would be useful to ask the students about their feelings in the communicative situation. How did they feel? Were they nervous or anxious about the upcoming test? What kind of day had they had so far? The chance of one having a bad day is something that is next to impossible to eliminate but asking these questions would give an insight on what factors could have resulted in a person being silent or talkative.

In addition, as this thesis is about the interlocutor’s effect, the student’s feelings about the interlocutor’s role in his/her own performance would be worth studying. How did it feel that the interlocutor was talkative/silent? How did it affect the student’s performance according to his/her own opinion?

Trying to make oral language proficiency testing as valid, reliable and objective as possible is not a simple task. The problems and questions present in all testing also affect oral language testing. Furthermore, the assessment of spoken language is especially tricky as the result of the test does not depend only on the test taker but also on the rater and (in case of pair or group discussion) on an interlocutor or
interlocutors. Even though it is impossible to eliminate all the variables affecting the validity and objectivity of a test, it is important to gain understanding and knowledge of them. The more research and material is done on interaction in oral language testing situation, the more the raters are aware of it. The qualitative studies (e.g. O'Sullivan 1995, 2002; Porter and Shen 1991 and Berry 1997) conducted on the interlocutor’s effect on the outcome of spoken language test provide relevant information on the complex communicative situation present in the spoken language test.

However complex an issue, oral language proficiency testing will be of growing importance in language assessment in the future. Nobody can deny the importance of being able to speak foreign languages in today's international world and the direction the world is rapidly changing will set new demands for language education and assessment as well. Communication (both written and oral), interaction and cooperation are central means of surviving in today's world and it is important that these elements are being include in the studies of language learning and teaching. The information we gather on oral language proficiency assessment will provide us with tools for keeping language education up-to-date with the rest of the changing world.
8. References:


Appendix 1

Tehtävä 1. Arkitilanteita

Keskustele parisi kanssa mahdollisimman luontevasti. Ilmaise vuorosanojen asiapisältö kohdekielellä. Älä käännyä, vaan yritä saada itsesi ymmärrettyksi omin sanoin. Jos et tiedä jotain, älä juutu vaikeaan kohtaan vaan jatka eteenpäin ja puhu mahdollisimman paljon.

Nico/Anna viipyy luonasi kuukauden, jonka aikana käytte seuraavat kaksi keskustelua (numerot 2.1 ja 2.2). Vaihtakaa vuoroja niin, että kumpikin teistä on toisessa tilanteessa oma itsensä (S=sinä) ja toisessa vieraan (Nico/Anna) roolissa. Sopikaa roolijako ennen kuin alatte puhua.

2.1 Lentoasemalla

S Tervehdi ja kerro, kuka olet ja varmista, että tulija on Nico/Anna.
N/A Tervehdi ja esittäydy kohteliaasti ja sano terveisä perheeltäsi.
S Toivota tulija tervetulleeksi ja kysy jotakin matkasta.
N/A Vastaa kohteliaasti ja kysy, mistä löydät vessan.
S Sano, että se on samassa kerroksessa, noin 30 metriä eteenpäin ja sitten vasemmalle.
N/A Kiitä ja pahoittele, että toinen joutuu odottamaan.
S Sano, ettei se haittaa. Bussikin lähtee vasta 15 minuutin kuluttua.
N/A Pyydä, että toinen vahtii laukkujasi ja lupaa yrittää palata nopeasti.

2.2 Koti-ilta

N/A Tiedustele, mitä illalla tehdään.
S Kysy, sopiiko vieraallesi koti-ilta.
N/A Sano, että sinulle sopii hyvin.
S Ehdota, että laittaisitte yhdessä jotain syötävää.
N/A Suostu ja kerro, mitä osaisit tehdä ja kysy, kelpaako se.
S Kerro mielipiteesi kohteliaasti. Pidätkö, etkös pidä, miksi?
N/A Ehdota, että ruuan jälkeen katsottaisiin telkkaria/dvd:ää.
S Ehdota jotain suosikkiasi.
Tehtävä 2. Neuvottelu retkipäivästä

Suunnittelette yhdessä retkeä johonkin suosittuun paikkaan kotiseudullasi. Sopikaa yhdessä seuraavista asioista:

- mihin retki tehdään, mihin aikaan ja mistä lähdetään
- keitä lähtee mukaan
- miten pitkä matka on ja miten se tehdään (kävellen/bussilla/pyörillä)
- mitä kumpikin haluaa tehdä ja nähdä
- missä syödään ja mitä
- miten paljon rahaa otetaan mukaan ja mihin sitä kuluu
- milloin palataan takaisin
- mitä pitää muistaa / mitä ei saa unohtaa
Appendix 2

Discussion 1

Student 3 (S3) and student 4 (S4)

<S4> er hello i’m <NAME> you must be nico </S4>
<S3> yea i’ m nico er and greetings from my family </S3>
<S4> erm welcome to finland so how was your journey was the flight really long </S4>
<S3> mm yea it’s wa it was quite long er <P: 05> er excuse me but where could i find a bathroom [(xx)] </S3>
<S4> [er] it’s in the same floor about thirty meters forward and then to the left </S4>
<S3> thanks er sorry you have to wait </S3>
<S4> oh that’s alright the bus will only leave in fifteen minutes </S4>
<S3> could you watch my er baggage and i’ll try to get back fast </S3>
<S4> okay so what should we do tonight </S4>
<S3> er should we go to a , a home evening evening </S3>
<S4> er yea that sounds wonderful </S4>
<S3> could we make some food together </S3>
<S4> er yea sure i could do i could do pea soup if that’s alright </S4>
<S3> well okay , i like it <S4> [@@] </S4> [it’s good] and cheap </S3>
<S4> er yea er how about if after we’ve eaten we watch tv or dvd </S4>
<S3> yeah let’s watch star wars </S3>
<S4> @@ </S4>

<S4> er , <FOREIGN> nyt mä oon niin <FOREIGN> er hey there’s a park nearby called kaivopuisto i was wondering if we could go there for a picnic it’s a wonderful weather </S4>
<S3> alright when should we go </S3>
<S4> er i don’t i was thinking tomorrow because it’s starting to get pretty late er at what time do you think you’d be up how long do you sleep </S4>
<S3> er quite long let’s say one am </S3>
<S4> okay wonderful so we’ll just leave at from our place , er i don’t any of my friends are coming you don’t know anyone no </S4>
<S3> no @@ </S3>
<S4> no er </S4>
<S3> er alright the journey isn’t long right </S3>
<S4> no it’s not it’s er just er , short (xx) we’ll just go there by tram and then we’ll walk </S4>
<S3> okay , well what would you like to do there </S3>
<S4> well i thought i could just eat and watch people play or birds flying </S4>
<S3> alright. what do we eat and where </S3>
<S4> erm i was thinking we could er both bring something maybe and then we could buy something from over there maybe like ice-cream ‘cause that could just melt in our bags </S4>
<S3> okay . [could we] <S4> [er] </S4> take money </S3>
yea sure, i think we’ll only buy ice-cream but if you wanna buy something else too then that’s just fine. at what time do you think we should come back

er some when we are ready

yeah, just remember to take the money with you ‘cause [okay] [picnics] really suck if you don’t eat any ice-cream
Appendix 3

Discussion 3

Student 5 (S5) and student 3 (S3)

<S5> hi my name is <NAME> er are you nico </S5>
<S3> yes i’m nico hi greetings from my family </S3>
<S5> erm welcome to finland erm how was your er trip </S5>
<S3> it was comfortable er where can i find a bathroom </S3>
<S5> er at the same floor about thirty meters er and then left </S5>
<S3> thanks sorry you’ll have to wait </S3>
<S5> er no it’s okay the bus leaves er five minutes ago </S5>
<S3> er could you watch my bags i’ll be back soon </S3>
<S5> yea </S5>
<S3> what could we do in the evening </S3>
<S5> maybe we can er er have a homenight </S5>
<S3> it’s fine to me </S3>
<S5> mhm , maybe we can cook something together </S5>
<S3> alright i could do spaghetti is it okay </S3>
<S5> yea i like spaghetti . erm </S5>
<S3> could we watch some tv or movies </S3>
<S5> yea , erm i like . erm funny movies </S5>
<S3> okay . alright , could we go to suomenlinna this afternoon </S3>
<S5> yea it’s good idea erm </S5>
<S3> we could go there at five o’clock from the harbour </S3>
<S5> yea , erm em can i take some my friends </S5>
<S3> okay , we could of course we have to the ferry there , what would you like to do and see there </S3>
<S5> i don’t know , we can eat something good and just walk around </S5>
<S3> alright there must be some kind of restaurant there </S3>
<S5> yea </S5>
<S3> let’s take something from there </S3>
<S5> yea </S5>
<S3> alright how much money should we take with us </S3>
<S5> maybe twenty euros or something </S5>
<S3> i think it’s far too little it’s still (xx) paying well of . when could we come back from there </S3>
<S5> er well maybe </S5>
<S3> nine o’clock </S3>
<S5> yea something like that </S5>
<S3> alright let’s take money with us and let’s try (xx) there to be there at the time </S3>
<S5> yea </S5>
Appendix 4

Discussion 3

Student 4 (S4) and student 6 (S6)

<S6> okay er hi are you anna </S6>
<S4> yes i am </S4>
<S6> okay hi i’m <NAME> </S6>
<S4> hello , so er my mom says hello she liked your letter very much </S4>
<S6> oh thank you , so welcome to finland and did you have a good race </S6>
<S4> yes it was wonderful trip , er do you know where the bathroom is </S4>
<S6> oh yea it’s in this same floor er you have to go just 30 metres forward and then turn left </S6>
<S4> thank you and sorry you’ll have to wait </S4>
<S6> oh it’s okay the bus is leaving just in fifteen minutes </S6>
<S4> okay great er could you please watch my luggage er i’ll try to be as quick as i can </S4>
<S6> okay </S6>
<S4> erm what shall we do tonight </S4>
<S6> oh can we stay at home </S6>
<S4> yes that’ll be just fine </S4>
<S6> okay erm maybe we can , cook or something </S6>
<S4> yea that sounds wonderful er i could do pasta if that’s okay </S4>
<S6> oh yea that’s good erm , i love pasta </S6>
<S4> @@ how about it if after dinner we’d watch tv </S4>
<S6> oh that would be wonderful i love titanic @@ </S6>
<S4> okay </S4>
<S6> er what would we do today </S6>
<S4> i think we could go into a famous place in helsinki do you know anything where we could go </S4>
<S6> do you mean an amusement park </S6>
<S4> yea that would be great </S4>
<S6> oh yea </S6>
<S4> er , what kind of rides is there do you know </S4>
<S6> erm i don’t know </S6>
<S4> you don’t know have you been there </S4>
<S6> er , no i haven’t been there do you </S6>
<S4> @no@ </S4>
<S6> okay when we are going there </S6>
<S4> er i don’t know maybe noon would be good </S4>
<S6> okay er </S6>
<S4> er do you know how long the er trip is </S4>
<S6> er about [twenty minutes] </S6>
<S4> [how long will it take] okay so shall we walk </S4>
<S6> er maybe or then we can take a bus </S6>
<S4> okay [we can] </S4>
<S6> [you can just] </S6> see if the bus comes in time for us
<S4> </S4>
<S6> so where we er where can we eat </S6>
well i was thinking if we’d eat at the amusement park, i suppose they have some kind of food stand [maybe french fries] there yea. i think i’ll take like twenty euros, do you think that would be enough

er maybe, that’s good

yea that’s right, great

what time we will come back to home

well we’ll see when we’re done

okay

okay, don’t forget to take your hat with you @@

@no i won’t
Appendix 5

Discussion 4

Student 5 (S5) and student 6 (S6)

<S6> hello i’m <NAME> are you anna </S6>
<S5> yea my name is anna and , erm i’m from . england </S5>
<S6> okay er welcome er , er did you have a good trip </S6>
<S5> yea er but erm where is the vc </S5>
<S6> er it’s in this same floor , you have to go just , 30 meters straight and then left </S6>
<S5> oo thank you sorry that you have to wait </S5>
<S6> it’s okay we have time before the bus is leaving </S6>
<S5> erm , er can you erm look my , bags </S5>
<S6> [erm] [erm] i came back quickly </S6>
<S6> okay , @@ okay . er what shall we do in the night </S6>
<S5> maybe we can er , make a home night </S5>
<S6> okay it’s sounds good </S6>
<S5> erm maybe we can cook something good </S5>
<S6> oh yes i ca i can cook i like cooking and , is it okay </S6>
<S5> yea er erm . i like cooking too </S5>
<S6> @@
<S6> (xx) er er er can we er watch @@ t t @@ can we watch tv before the dinner </S6>
<S5> @i like salkkarit@</S5>
<S6> @okay@ <FOREIGN>voi ei noni sit tota noi<FOREIGN> er , where are we going today </S6>
<S5> maybe we can go to korkeasaari i like animals </S5>
<S6> @@ er okay from where are we leaving to korkeasaari </S6>
<S5> erm i don’t know , maybe we can go by bus and we can , er , go there . er at er three o’clock </S5>
<S6> okay . er so is it okay that my brother is coming [with] </S6>
<S5> @ya] </S5> us </S6>
<S5> what would you like to see there </S5>
<S6> oh i love animals and i want to see them and everything </S6>
<S5> yea , and there is very good ice-cream and sandwiches we can eat </S5>
<S6> oh yes that sounds good , er how much money we need </S6>
<S5> erm maybe something five euros or , something </S5>
<S6> okay , er so when we are leaving </S6>
<S5> maybe at three o’clock </S5>
<S6> okay that’s good