THE IMPACT OF TASK TYPE ON ORAL PERFORMANCE OF ENGLISH LANGUAGE PREPARATORY SCHOOL STUDENTS

GÖREV TÜRLERİNİN İNGİLİZCE HAZIRLIK OKULU ÖĞRENCİLERİNİN SÖZLÜ PERFORMANSLARI ÜZERİNE ETKİSİ

Meltem Huri BATURAY [*], Hatice SANCAR TOKMAK[**], Berrin DOGUSOY [***], Aysegul DALOGLU [****]

ABSTRACT: In this study, the effects of narrative, descriptive and prediction-personal reaction task types with visuals on the oral performance of intermediate level English language learners were compared. The study was carried out at Gazi University Preparatory School, Research and Application Center for the Instruction of Foreign Languages with the participation of eighty five students. The students' oral exam marks based on the assessment of narrative, descriptive and prediction-personal reaction task types for each group were compared and focus-group interviews were carried out with the students in order to get in depth data on the issue. The results showed that there was no significant difference between the narrative, descriptive and prediction-personal reaction task type groups in terms of oral performance scores. However, qualitative data showed that the narrative group students felt more relaxed and free to express themselves in the assessment sessions.

Keywords: Speaking, Oral Assessment, Task Types

ÖZET: Bu çalışmada, görsellerle sunulan anlatımsal, tanımsal ve tahminsel-kişisel reaksiyon task türlerinin orta seviye İngilizce hazırlık sınıfı öğrencilerinin oral performansları üzerindeki etkisi karşılaştırılmıştır. Çalışma Gazi Üniversitesi Yabancı Diller Öğretimi Araştırma ve Uygulama Merkezinde, 85 öğrencinin katılımıyla gerçekleştirilmiştir.. Anlatımsal, tanımsal ve tahminsel-kişisel reaksiyon task türlerinden birine göre değerlendirilen öğrencilerin sözlü sınav notları karşılaştırılmış ve konu ile ilgili daha derinlemesine bilgi toplamak amacıyla odak görüşme yöntemi uygulanmıştır. Sonuçlar, anlatımsal, tanımlayıcı ve tahminsel-kişisel reaksiyon gruplar arasında sözlü performans notları açısından önemli bir fark olmadığını göstermiştir. Fakat nitel veri, anlatımsal grup öğrencilerinin kendilerini değerlendirmeler sırasında daha sakin olduklarını ve kendilerini ifade etmekte konusunda daha özgür hissettiklerini ortaya koymuştur.

Anahtar sözcükler: Konuşma, Oral Değerlendirme, Task Türleri

1. INTRODUCTION

Today, communication skills are taught in a wide range of general education courses and students are in need of speaking and listening skills that will help them succeed in future courses and in the workplace. Thus, the assessment of communication skills is an important issue in general education (Dunbar, Brooks and Miller, 2006). Oral assessment is often carried out to look for students' ability to produce words and phrases by evaluating students' fulfillment of a variety of tasks such as asking and answering questions about themselves, doing role-plays, making up mini-dialogues, defining or talking about some pictures or talking about given theme. As categorized by Bygate (1987) the operations in an oral ability test are either informational skills or interactional skills.

Nakamura (1993) stated that testing oral proficiency became an important issue with the emergence of communicative language teaching in which speaking skill had a prominent role. Regarding speaking skill, Madsen (1983) declared that "The testing of speaking is widely regarded as the most challenging of all language tests to prepare, administer and score" (p. 147). The reasons for the difficulty of the assessment of speaking ability or the oral exams emanates from, firstly, the nature of the speaking skill since it is not easy to decide whether fluency or accuracy will be evaluated and

[*] Yrd. Doç.Dr., Kırıkkale Üniversitesi, baturay@kku.edu.tr
[**] Dr., Mersin Üniversitesi, haticesancarr@gmail.com
[***] Arş. Gör., Mersin Üniversitesi, bdogusoy@metu.edu.tr
[****] Doç. Dr., Orta Doğu Teknik Üniversitesi, daloglu@metu.edu.tr

the criteria to evaluate the performance of the exam takers; secondly, the role that the tester will play during an oral assessment has to be decided prior to the assessment. Hingle and Linington (2002) stated that people whom are involved stand in the fore front in oral assessments, more than the testing instrument. In addition to these, task demands and task support are two important concepts that might change students' achievement and interest in the oral assessment test. As Taguchi (2007) confirmed, features of second language oral output such as accuracy, fluency and complexity vary by task type.

1.1. Picture Use in Oral Assessment

As Peng and Levin (1979) and Anglin (1986) referred to, pictures are easy-to-recall memories and it is easy to store them in the long term memory. It was found in some research studies that one text is remembered better when it is illustrated with pictures (Ollerenshaw, Aidman and Kidd, 1997; Schnotz, 2002). As stated by Kennedy (1974), pictures are often fairly precise and unambiguous in their reference of objects; therefore, usually all people see the same thing in a picture. For that reason, pictures are believed to be ideally suitable for eliciting oral performance of the students. As stated by Crisp and Sweiry (2006), there is not much research into the effects of including visual resources in examination questions; therefore, the research on the influences of illustrations in instructional texts might provide some relevant insights. However, it is essential to carefully select the pictures which are going to be used in an oral assessment. Some pictures might be attractive, but may not lend themselves to a variety of language. Also, when a picture-based oral assessment is the issue, mostly cognitive, language and metalinguistic demands and support exist within the tasks (Cameron, 2001).

1.2.Task Types Used in the Study

1.2.1. Narrative Tasks

The provision of a general context in the narrative task further includes language support by enabling the student use of words and phrases already known. The task does not limit but enables the student to apply whatever language s/he possesses. As Skehan and Foster (1999) emphasize a narrative task which is based on a cartoon strip, with the provision of a clear inherent structure particularly in terms of time sequence helps to ease the processing burden of the task and leads to more fluent and accurate performance.

In this study, the researchers named the oral assessment task as 'narrative task' which required students to produce oral statements by examining the pictures. The only instruction in this task type was 'tell the story in the pictures'.

1.2.2. Descriptive Tasks

A descriptive task has been defined by the authors of this paper as the task in which students are required to answer the questions accompanied with the pictures. Because there are some questions to be answered, students are expected to find the clues in the pictures to reach the acceptable answers.

In this study, the descriptive task students were given pictures depicting a scene followed by questions such as 'how many people are there in the picture?', 'can you describe the people in the picture?' 'what are they doing?,

1.2.3. Prediction-Personal Reaction Tasks

As in the descriptive task, the prediction- personal reaction task contextualizes the events accompanied by some questions about the details depicted in the pictures; however, it differs from the descriptive task in that way, students are expected to make predictions and utter some predictive statements regarding the pictures. Like narrative task group, the students are not bounded by questions and can focus on the narrative side of the pictures. For that reason, this task type has similar features to both descriptive and narrative task types. Therefore, it was categorized as a different task type from descriptive and narrative ones. Because this task type requires 'differentiated outcomes' as defined by Skehan (1998a; 1998b) the complexity of the language increases (Fulcher and Marquez Reiter, 2003).

1.3. Significance of the Study

Oral assessment is a widely used for testing instrument and task types have an affect on second language oral output such as accuracy, fluency (Taguchi, 2007). Also, the research studies investigating the effects of including visual resources in examination questions are limited in the literature (Crisp and Sweiry, 2006). Since there is a gap in the literature in terms of effects of visual resources in assessment, , research studies which focus on the effects of different there task types presented with visuals on students' oral performances are required .

Moreover, students' perceptions about evaluation methods play a significant role since their perceptions about assessment significantly influence their approaches to learning and studying (Struyven, Dochy and Janssens, 2005). This study had interesting results in terms of the perceptions of the students regarding oral assessment.

2. METHODOLOGY

This study focused on the range of oral assessment task types with the aim of comparing the effects of narrative, descriptive and prediction-personal reaction task types with a focus on their task demands and support on the oral performance of intermediate level English language students. More specifically, the study aimed to answer the following research questions:

1. Do students' oral performance scores on narrative, descriptive and prediction-personal reaction tasks differ?
2. What are the students' perceptions regarding the task demands of narrative, descriptive and prediction-personal reaction oral assessment tasks?

2.1. Participants

85 preparatory school students in four intermediate level English language classes participated in the study. All of the participants were preparatory school students who were in their first year at university and the university enforced them to take a one-year intensive English course at prep school because some courses in their own departments would be in English the following year. At the beginning of the semester, students took the English proficiency test prepared by the school and according to the exam results they were classified into levels. Depending on this classification, all the students in this study were thought to be at the same level of English proficiency (B level) and oral skill capability. There were 21 students in class B1, 21 in class B2, 22 in class B3 and 21 in class B4. The oral assessment sessions of these groups were carried out by four teachers in groups of two. The classes B1 and B2 were assessed by the first pair of teachers and B3 and B4 were assessed by the second pair of teachers. The participants of the study were between 19 and 25 years old. The participants of the focus group interview were chosen randomly by the researchers.

2.2. Procedure

The purpose of the oral assessment sessions at Gazi University Preparatory School was to identify students' oral ability in English by using an assessment criteria which focused on pronunciation, use of vocabulary and grammar, fluency, and thematic relevance which required students to tailor their speech in a sociolinguistic appropriate manner to the individuals and circumstances presented in the given pictures and to the given themes. The students were required to talk about given pictures or a set of illustrations by narrating or describing them. According to Bygate's category (1987), the students were asked to describe sequence of events, make comparisons, give explanations, express opinions, draw conclusions and make comments. Students were randomly given the tasks and thus they fell into the groups of descriptive, narrative or prediction-personal reaction task groups in the study.

All examiners were guided to follow the same procedure. Every instructor was guided to assess a class other than his/her own class and they worked in pairs. As suggested by Hughes (2003), the

second examiners helped to the first ones to keep track of the exam takers' performance and to reach a reliable judgment on the range of scales. Prior to oral assessment, every instructor was provided with clear written instructions about the procedure.

In order to ensure valid and reliable scoring, every instructor was informed about the criteria for the rating scale which included the skills of grammar and vocabulary use, pronunciation accuracy, fluency and finally the thematic relevance. Instructors were expected to evaluate the students' performance by giving a score between 0 (zero) and 5 (five) for each criteria. The assessment sessions took approximately 5-7 minutes and there was not a time limit.

Almost two weeks after their oral assessment exam, 12 students were interviewed together by using focus group method. The researchers were able to get detailed information which task type the students favored and why.

2.3. Instruments for Data Collection

2.3.1. Pictures

It is believed that pictures are ideal materials can be used to elicit oral performance of the students. However, in a picture-based oral assessment exam, they may help or limit exam takers. Also, in these types of assessments, there are mostly cognitive, language and metalinguistic demands. For that reason, the pictures should be selected carefully. As mentioned before, in this study three tasks were used in the oral assessment. Four experts selected the potential pictures which would be used in the oral assessment, and then the pictures they agreed were used in the study.

2.3.2. Focus Group Interviews

In the study one of the data collection methods was focus group interview and it was conducted almost two weeks after the oral assessment exam. According to Morgan (1997), this is both an important and a beneficial method when it is combined with the participant observation. In addition to this, cost effectiveness is another important side of this method as it enables the researcher to interview a number of participants at one time. The researchers' main aim was to create a supportive atmosphere in which discussion was promoted and to give chance to the participants to explain their points of views with marked composure. For that reason, the researchers asked more general questions such as "What do you think about the assessment method?"; "What do you think about the oral assessment type you were involved?"; "Do you think you could be more successful if you were in the other group?"

According to Marshall and Rossman (1999), focus group interviewing generally includes 7 to 10 people, but the size of the group can change. Smaller groups can consist of four people and larger groups can include twelve people who do not know each other. In this study, 12 students participated in the focus group interview.

The focus group interview method has some disadvantages such as limited control of the researcher over the group compared to one-to-one interviews (Krueger, 1988), data is more difficult to analyze, and it is hard to handle the group discussion when there is an emerging irrelevant discussion during the interview (Marshall and Rossman, 1999). In the study, the focus group interview results were transcribed by two researchers to provide the intercoder reliability which means high consistency among different coders. This is implemented by having someone else code the responses again (Johnson and Christensen, 2004). There was 90% agreement between the researchers on the themes emerged from the data.

2.3.3. Evaluation Rubric

Evaluation rubric was prepared by the teachers of the preparatory school who were also assessors in the study. The rubric was consisted of five analytic criteria as grammar and vocabulary use, pronunciation, fluency and finally thematic relevance some of which were suggested in Canadian

Language Benchmarks 2000: Theoretical Framework (Pawlikowska- Smith, 2002). All the students were given a score between 0 (zero) and 5 (five) for each category according to their performances.

2.4. Data analysis

A one way - ANOVA was performed to investigate the effects of the narrative, descriptive and prediction-personal reaction task types on the students' oral assessment performance. To see if there was a difference between the narrative, descriptive and prediction-personal reaction tasks, scores from grammar and vocabulary use, pronunciation accuracy, fluency and thematic relevance categories, a MANOVA analysis was further computed. On the other hand, data regarding students' perceptions on narrative, descriptive and prediction-personal reaction tasks types were collected through focus group interviews and categorized qualitatively.

3. FINDINGS

In this part, the findings of the study were reported under the themes of teacher difference, task types and students' perceptions on different task types. Teacher difference was also taken into account in that the examiners' difference may affect the performance of the students. For that reason, this treatment effect was additionally analyzed by using a one-way ANOVA.

3.1. Teacher Difference

It was important to acknowledge if there was a difference between the groups which were assessed by the pair of teachers 1 and in the study 2 (the classes B1 and B2 were assessed by the first pair of teachers and B3 and B4 were assessed by the second pair of teachers) in order to increase the reliability of the results. The same procedure was applied by the different teachers and they made the assessment according to a rubric which consisted of the skills of criteria grammar, vocabulary use, pronunciation accuracy, fluency and thematic relevance. In order to see whether there was a difference among students' performances which might be caused by different teacher affect, a one-way ANOVA was performed. The results indicated that teacher difference had no affect on students' oral assessment performance with $F(1,83)= 3.43$, $P>0.05$. All students' oral performance scores for the first pair of teachers was M=5.92 with SD= 1.28 while for pair two it was M=6.44 with SD=1.30. Although the mean scores indicated that the group of second pair of teachers seemed to get higher scores than the group of first pair of teachers group, SD for the second pair of teachers group was higher than the other one.

3.2. Task Type Difference

Although examiners were informed about the procedure prior to the oral assessment in order to ensure all exam takers have an equal opportunity to display their abilities, and the same assessment items were used for all classes, as suggested by Alderson, Clapham and Wall (2001), some of the exam takers were only required to make simple but appropriate comments while the others were forced to use complex language. This was due to different task types and different task demands that resulted from these different tasks.

The research question one, "Do students' oral performance scores on narrative, descriptive and prediction-personal reaction tasks differ?" was analyzed with a one-way ANOVA. The number of the students assessed with the narrative task was 29, the number of the students assessed with descriptive task was 28 and the number of the students assessed with prediction-personal task type group was 28. According to the analysis results, the students' oral assessment scores did not differ in the narrative, descriptive and prediction-personal reaction tasks types with $F(2,82)=1.68$, $P>0.05$ (Table 1). In addition, a MANOVA was performed to see if the students' scores they got from each categories as "grammar and vocabulary use", "pronunciation accuracy", "fluency" and finally "thematic relevance" differ in terms of narrative, descriptive and prediction-personal reaction task types. According to the outcomes of the MANOVA analysis, task types did not seem to affect the students' "grammar use" score with $F(2,82)=2.32$, $P>0.05$; "vocabulary use" score with $F(2,82)=0.73$, $P>0.05$; "pronunciation

accuracy" score with $F(2,82)=0.71$, $P>0.05$, "fluency" with $F(2,82)=1.11$, $P>0.05$ and "thematic relevance" score $F(2,82)=0.58$, $P>0.05$.

3.3. Students' Perception on Narrative, Descriptive and Prediction-Personal reaction Task Types

The data collected in the focus group interviews aimed to answer the second research question. Of the twelve volunteers were participated to the focus group interview. Regarding the task type they had in the oral exam, two of the participants took the narrative task, four took the descriptive task, and six took the prediction–personal reaction task. The collected data by using focus group interviews were coded by the two researchers and emerged themes agreement between them was very high at 90%. The emerged themes were "Accordance with the task type", "Limitations of the task type", "Emotional Interpretations", and "Difficulty in use of imagination". These emerged themes were placed under the each group types as Narrative, Descriptive and Prediction-Personal.

When participants were asked which task they preferred for their oral performance assessments, they favored narrative task more than the others and they started to compare the narrative task with the other two tasks (descriptive and prediction-personal reaction). The participants who completed the narrative task stated that they felt more comfortable during the speaking activity while other participants stated that they felt more bounded to the questions which were asked by the examiners.

3.3.1. Narrative Task Group

3.3.1.1. Accordance with the Task Type

The codes under this theme were "not bounded to the questions", "convenience of the questions", "flexibility" and "comfortable". The students of the narrative task group pointed out that the task was more advantageous than the other tasks in that they were not restricted by the examiners. The examiners just told them to speak on the pictures showed. They did not have to use a specific tense or a specific grammar rule. The assessment was speech-based and for that reason, they could freely speak on the pictures showed. As Omaggio (1979) emphasized students faced with a less demanding task with the provision of the contextual information that the narrative held. The narrative task group additionally stated that they felt themselves more comfortable while explaining what they had seen in the pictures because they were not bounded with any questions. Also, they pointed out that they were far away from tension and stress because the examiners did not expect them to use any specific grammar rule or vocabulary. One participant in Narrative group states that:

> *"...I feel free myself in terms of speaking. During the exam, if I was stuck in one scene of the pictures, I could skip to the other scene and used the grammar structure whatever I wanted..."*

3.3.2. Descriptive Task Group

3.3.2.1. Accordance with the Task Type

Only one participant from the descriptive task group stated that she had not had any problems and difficulty regarding the task while speaking. She added that it did not matter for her whether the task was a narrative one or any other. However, when the task that was performed by this participant in the oral assessment was investigated it was observed that her question had a more narrative feature although it had directions. The question which was asked to this participant was, 'the girl in the left corner is Nicola. Look at the room and talk about her hobbies/ leisure time activities and daily routines.'

3.3.2.2. Limitations of Task Type

The codes under this theme were "grammar-based contraction" and "stepwise structure". Three of the four participants who completed the descriptive task stated that they could have been more successful and could have spoken more fluently if they had been in the narrative task group. Also, they confirmed that narrating the pictures was easier than answering the questions about the pictures.

3.3.2.3. Emotional Interpretations

The codes under this theme were "bounded to the questions" and "inflexibility", "anxiety". The descriptive group students stated that felt uncomfortable and more bounded to the questions and they were and more nervous in contrast to the other task groups due to the constraint of using the same grammar structure which the questions were constructed by. One of the participants stated that;

"...you could find more things to say while narrating the pictures...There was a girl in my picture and examiners said that she had a lot of work to do. According to the picture, examiners asked me to describe firstly, the house in the picture and then talk about what work she should do. I could speak only for a few seconds because I felt that I was only required to make up sentences with 'should' structure as in the questions..."

3.3.3. Prediction-Personal Task Group

3.3.3.1. Accordance with the Task Type

The prediction-personal task group participants made comments similar to the descriptive task group in that they stated that they could have been more successful if they had been assessed with a narrative task because they were directed by the examiners' questions.

3.3.3.2. Limitations of Task Type

The codes under this theme were "grammar-based construction" and "bounded to the questions and pictures". The prediction-personal task group participants stated that they focused on the specific grammar rules. For example, they had to use future tense or past tense according to the questions. There are two participants' statements on these issues:

Similar to descriptive task group, they felt themselves bounded to the questions; they just answered the questions and did not have the opportunity to add any extra words to their speeches. One of the participants stated that:

"....it was the worst speaking performance for me. The questions restricted me..."

3.3.3.3. Difficulty in Use of Imagination

Since the prediction-personal task type required a predictive statement in itself which was also related with the cognitive processes of the individuals which is also connected with the linguistic ability, they found this task more difficult than the other ones particularly while performing it in English because students felt themselves stuck while making predictive statements and they had difficulties in finding what they had to say about the pictures. As it required use of grammatical structure of prediction and related vocabulary, it increased the cognitive demand of the task.

4. CONCLUSION AND DISCUSSION

This study aimed to investigate the effects of the different task types group as narrative, descriptive and prediction-personal reaction on English Preparatory School students' oral performance. The sampling of the study consisted of 85 English Preparatory School students form Gazi University.

Students' trust towards the testing procedure is very much related with the tasks they are required to perform. In such circumstances, the exam takers should question whether the assessment measures the achievement of the learning outcomes in a reliable and fair manner for all exam takers. In this study, the main factor that was believed to affect the exam takers' trust towards the testing procedure was the teacher difference. Therefore, in the study the teachers used the same procedures and rubric for the assessment of the students. Also, each group was assessed by two teachers and their agreement was the first requisite for this study. The qualitative validity arguments were supported with the statistical analysis in order to investigate whether there was a difference between the group scores

in terms of teacher difference. The ANOVA results showed that there was no significant difference between the students' oral assessment scores in terms of the teacher difference.

Also, the ANOVA results showed that there was no significant difference between the students' oral assessment scores in terms of task types. However, it was found out that the students had different perceptions towards the oral assessment tasks. According to the qualitative data results, students pointed out that they felt restricted with the questions in the descriptive task and had answered just the questions without adding extra comments to the events depicted in the pictures. Similarly, the prediction-personal task group found the questions more complex because of the characteristic of the prediction concept itself. Since making predictions according to the pictures required a cognitive process and linguistic ability, the students stated that they had difficulty in predicting events depicted in the pictures. Besides, making prediction statement required some metalinguistic knowledge about the language in production. Moreover, as Skehan (1998a; 1998b) pointed out the psycholinguistic category "differentiated outcome" deals with the more complex outcomes there was only one question as "Tell the story in the picture". This advantage of the narrative task might again be explained again with Skehan's (1998a; 1998b) category "structured task" which lead to significantly greater fluency and accuracy. Students in this group stated that they felt themselves freer to speak about the pictures although the results of the study which was focused on quantitative data indicated that there was no significant difference between the student's oral assessment scores in terms of task types, namely, descriptive, narrative and prediction-personal.

Besides the task, culture, on the other hand, is another factor affecting receiver's perception of the message in the visuals. There are various visual representations of the message such as signs, symbols, images and illustrations, graphics, diagrams, charts. It is obvious that two persons with the same age but from different cultures will perceive a visual message in different ways. Thus, the difference in their perceptions might affect their performance in carrying out the tasks, particularly at the descriptive task in which students are required to answer the questions. The reasons for the non-significant difference of the quantitative results of the study may be due to the background of the students, the students' competency differences in English and the characteristics of the visuals used for each task type. Similar studies should be performed to see if these results are replicated. The researchers might take into account the effects of the culture, visuals and students' competency in English so they might find different relations between task types and the students' performance. To recap, it is hoped that future oral assessments with the visuals will be designed to undertake the findings of the present research.

To recap, although the statistical analysis indicated no significant difference on the oral achievements of the students in terms of task performance difference, the perceptions of the students pointed out that tasks have different burdens on the students. It is obvious that, some tasks overburden students due to the demands they require but some facilitate the work of the learners with the support they provide for them. Therefore, demand and the support of the oral assessment tasks should always be balanced in order to provide equal opportunities and a fair assessment atmosphere for students.

REFERENCES

Alderson, J., Clapham, C., & Wall, D. (2001). Language test construction and evaluation. Cambridge: Cambridge University Press.

Anglin, G. J. (1986). Effect of pictures on recall of written prose: How durable are picture effects? Paper presented at the Annual convention of the Association for Educational Communications and Technology, Las Vegas, NV. (ERIC Document Reproduction Service No. ED 267 755)

Bygate, M. (1987). Speaking. Oxford: Oxford University Press.

Cameron, L. (2001). Teaching Languages to Young Learners. Cambridge: Cambridge University Press.

Johnson, B. & Christensen, L. (2004). Educational Research: Quantitative, Qualitative, and Mixed Approaches. USA: Pearson Education, Inc.

Crisp, V. & Sweiry, E. (2006). Can a picture ruin a thousand words? The effects of visual resources in exam questions. Educational Research, 48(2): 139 – 154.

Dunbar, N.E., Brooks, C.F. & Miller, T. K., (2006). Oral communication skills in higher education: using a performance-based evaluation rubric to assess communication skills. Innovative Higher Education, 31(2): 115-128.

Fulcher, G. & Marquez Reiter, R. (2003). Task difficulty in speaking tests. Language Testing, 20, 321-344.

Hingle, I. & Linington, V. (2002) English proficiency test: the oral component of a primary school. In Richards J.C. and Renandya, W.A. (2002), Eds. Methodology in Language Teaching. An Anthology of Current Practice. Cambridge: Cambridge University Press.

Hughes, A. (2003). Testing for Language Teachers (2nd Edition). Cambridge: Cambridge University Press.

Kennedy, J. M., (1974) A psychology of picture perception: Images and information. San Francisco: Jossey-Bass.

Krueger, R. A. (1988). Focus groups: A practical guide for applied research. Newbury Park, CA: Sage Publications.

Madsen, H. S. (1983). Techniques in testing. New York: Oxford University Press.

Marshall,C. & Rossman,G.B. (1999). Designing Qualitative Research (3rd Edition). Thousand Oaks: Sage.

Morgan, D. L. (1997). Focus groups as qualitative research (2nd Ed.). Thousand Oaks, CA: Sage.

Nakamura, Y. (1993). Measurement of Japanese college students' English Speaking ability in a classroom setting. Unpublished doctoral dissertation, International Christian University, Tokyo.

Ollerenshaw, A, Aidman, E & Kidd, G. (1997) Is an illustration always worth ten thousand words? Effects of prior knowledge, learning style and multimedia illustrations on text comprehension. International Journal of Instructional Media, 24(3): 227 – 238.

Omaggio, A.C. (1979). Pictures and second language comprehension: Do they help? Foreign Language Annals, 12(2): 107 -116.

Pawlikowska- Smith, G. (2002). Canadian language benchmarks 2000: Theoretical framework. Retrieved June 27, 2008 from http://www.language.ca/pdfs/final_theoreticalframework3.pdf

Peng, C-Y. & Levin, J. R. (1979). Pictures and children's story recall: Some questions of durability. Educational Communication and Technology Journal, 27(1): 39-44.

Schnotz, W. (2002). Towards an integrated view of learning from text and visual displays, Educational Psychology Review, 14(1): 101 – 120.

Skehan, P. and Foster, P. (1999), The influence of task structure and processing conditions on narrative retellings. Language Learning 49(1): 93-120.

Skehan, P. (1998a). A cognitive approach to language learning. Oxford: Oxford University Press.

Skehan, P. (1998b). Processing perspectives to second language development, instruction, performance and assessment. Thames Valley Working Papers in Applied Linguistics 4, 70-88.

Struyven K., Dochy F. & Janssens F. (2005). Students' perceptions about evaluation and assessment in higher education: a review. Assessment & Evaluation in Higher Education, 30 (4): 331-347.

Taguchi, N. (2007). Task difficulty in oral speech act production. Applied Linguistics, 28(1): 113 - 135.

Genişletilmiş Özet

Öğrencilerin sözlü performanslarının değerlendirilmesi günümüzde kullanımı gittikçe yaygınlaşan bir değerlendirme yöntemidir. Ancak, görev türleri öğrencilerin doğru ve akıcı konuşmalarını etkilemektedir (Taguchi, 2007). Bunun yanı sıra, Crisp and Sweiry (2006) belirttikleri gibi görsellerin sınavlarda kullanımlarının öğrencilerin performansına etkileri ile ilgili sınırlı sayıda çalışma bulunmaktadır. Bu nedenle bu çalışmada, resimlerle birlikte sunulan anlatımsal, tanımsal, tahminsel-kişisel reaksiyon görev türlerindeki soruların öğrenenlerin sözlü performanslarına etkisi telaffuz, kelime kullanımı, dilbilgisi, akıcılık ve konuya aitlik açılarından karşılaştırılmıştır. Anlatımsal görev türü sorularda öğrencilere sadece resimde betimlenen hikâyeyi anlatmaları istenirken, tanımsal görev türünde öğrencilere resimlerle ilgili çeşitli sorular sorulmuş ve bu sorulara cevap vermeleri istenmiştir. Tahminsel-kişisel reaksiyon türü sorularda ise öğrencilerden resimde betimlenen olaylarla ilgili yorumlar yapmaları istenmiştir.

Çalışmaya 85 öğrenci katılmıştır ve bu öğrencilerin hepsi Gazi Üniversitesi İngilizce Hazırlık Okulunda bulunan orta seviye (intermediate level) sınıf (B1, B2, B3 ve B4) öğrencilerinden oluşmuştur. Hughes (2003), sınavlarda iki değerlendirmecinin yer almasının daha güvenilir notlandırma yapılabilmesi açısından önemli olduğunu belirtmiştir. Bu öneriye bağlı kalınarak çalışmada daha güvenilir notlandırma yapılabilmesi amacı ile değerlendirmeler, 4 öğretmen tarafından 2'erli gruplar şeklinde yapılmıştır. Diğer bir ifade ile B1 ve B2 sınıfında bulunan 42 öğrenci 2 öğretmen tarafından değerlendirilirken, B3 ve B4 sınıflarında bulunan 43 öğrenci diğer 2 İngilizce öğretmeni tarafından değerlendirilmiştir. Öğrencilerin 29'una anlatımsal görev türlerinden oluşan sorular sorulurken, 28'ine tanımsal görev türlerinden oluşan sorular, diğer 28'ine de tahminsel-kişisel reaksiyon task türü sorular sorulmuştur. B1, B2, B3 ve B4 sınıfı öğrencilerine rastgele olarak sorulan

task türleri değerlendirmeyi yapan öğretmenlerin ortak kararı ile belirlenmiştir. Çalışmada, öğrencilere herhangi bir zaman kısıtlaması yapılmamıştır ancak her bir öğrencinin değerlendirmesi ortalama olarak 5-7 dakika kadar sürmüştür. Değerlendirmeler sonrasında, anlatımsal, tanımsal ve tahminsel-kişisel reaksiyon görev türlerinden birine göre değerlendirilen öğrenci gruplarının sözlü performans notları karşılaştırılmıştır. Ayrıca, konu ile ilgili daha derinlemesine bilgi toplamak amacıyla odak görüşme yöntemi uygulanmıştır.

Değerlendirmecinin öğrenci üzerinde önemli bir etkisi vardır. Bu nedenle bu çalışmada değerlendirmecilerin etkisini azaltmak için öğretmenlerin uyguladıkları değerlendirme prosedürü ve değerlendirme formunun aynı olması sağlanmıştır. Bunun yanı sıra öğretmen faktörünün etkisi, farklı öğretmenler tarafından değerlendirilen öğrencilerin notları karşılaştırılarak istatistiksel veri analiz yöntemi kullanılarak incelenmiştir. Farklı öğretmenler tarafından değerlendirilen öğrencilerin notları arasında anlamlı bir fark bulunamamıştır. Başka bir ifade ile öğretmen farklılığının öğrenci performansı üzerinde herhangi bir etkisi yoktur.

Öğretmen farklılığının öğrencilerin sözlü performansları üzerine etkisinin olmaması, farklı 3 görev türünde sorular sorulan grupların performanslarının istatistiksel açıdan değerlendirilmesini ve grupların performansları arasında bulunabilecek herhangi bir farkın görev türlerinden kaynaklanabileceği yorumunun çıkarılmasına olanak sağlamıştır. Bu amaçla, 29 öğrenciden oluşan anlatımsal görev türü, 28 kişiden oluşan tanımsal görev türü ve 28 kişiden oluşan tahminsel-kişisel reaksiyon görev türü gruplarının sözlü performans notları istatistiksel analize sokulmuş ve gruplar arası önemli bir fark bulunamamıştır. Ayrıca, farklı görev gruplarının telaffuz, kelime kullanımı, dilbilgisi, akıcılık ve konuya aitlik notları ayrı ayrı analize sokularak aralarında bir fark olup olmadığı incelenmiş ancak anlamlı bir fark bulunamamıştır. Ancak, 12 gönüllü öğrenci ile yapılan odak görüşme yöntemi verilerine göre anlatımsal görev türü sorular yöneltilen öğrenciler kendilerini sınav sırasında oldukça rahat olduklarını ve kendilerini ifade etme konusunda özgür hissettiklerini belirtirken, tanımsal grup öğrencileri kendilerini sorularla kısıtlandırılmış hissettiklerini belirtmişlerdir. Ayrıca, tanımsal grup öğrencileri kendilerine sorularda yöneltilen soru zamanlarına ve kalıplarına bağlı kalmak zorunda kaldıklarından stres olduklarını bu nedenle de sadece kendilerine yöneltilen soruları kısa cümlelerle cevapladıklarını vurgulamışlardır. Benzer nedenlerle olmasa da tahminsel-kişisel reaksiyon grup öğrencileri değerlendirmeler sırasında kendilerini rahat hissetmediklerini belirtmişlerdir. Bu grup öğrencileri, değerlendirmeler sırasında istedikleri performansı gösterememelerinin nedenini sorulan soruların düşünme gerektiren sorular olmasından kaynaklandığını belirtmişlerdir. Ayrıca bu grup öğrencileri, resimde canlandırılan olayların tahmin edilmesine yönelik soruların kişiden kişiye değişiklik gösterebileceğini ve değerlendirmecilerin ne tür bir cevap aradıklarını bilmemelerinin kendileri üzerinde stres yarattığını vurgulamışlardır.

Özetle bu çalışmada istatiksel veri analizine göre 3 görev türü grubu arasında anlamlı bir fark bulunamamasına rağmen, odak görüşme yöntemi verilerinin analizine göre anlatımsal görev türü sorular sorulan grup değerlendirmeler sırasında kendini daha rahat hissettiklerini belirtmişlerdir. Çalışma sonuçlarından da anlaşılacağı üzere 3 görev türü öğrenciler üzerinde farklı etkiler yaratmıştır. Struyven, Dochy ve Janssens (2005) tarafından belirtildiği gibi öğrencilerin değerlendirme metotları üzerindeki düşünceleri onların öğrenme ve çalışma yöntemlerini de etkilemektedirler. Bu nedenle sınavlarda kullanılabilecek bu görev türlerinin değerlendirmecilerin amaçlarına göre seçilmesi önem taşımaktadır. Örneğin; değerlendirmeciler öğrencilerin konuşmaları sırasında dilbilgisi yönünden yanlışlıklar yaptıklarını düşünüyorlarsa sınavlarda tanımsal task türü soruları tercih ederken, öğrencilerin akıcı bir şekilde konuşmakta zorluk çektiklerini düşünüyorlarsa anlatımsal ya da tahminsel-kişisel reaksiyon task türü soruları tercih etmelidirler.

Bu çalışmanın sınırlılığı görev türlerinin farklı resimlerle sunulmasıdır. Cameron (2001) tarafından belirtildiği gibi resimlerin yapısı öğrencilerin performanslarını etkileyebilir. Bu nedenle, çalışmada resimlerin farklılığından kaynaklanabilecek etkiyi azaltmak için öğrencilere sunulan resimler ve bu resimlerle hangi görev türünün sorulacağı, konusunda uzman 4 öğretmenin anlaşması sonucu belirlenmiştir. Ancak bu etkinin tamamen sıfıra indirgenmesi amacıyla, ileride yapılacak çalışmada 3 görev türünde soruların sorulabileceği bir resim seçilerek bu çalışmanın tekrar edilmesi düşünülmektedir.