



THE IMPACT OF CONTEXT AND CONTENT VISUALS IN MULTI-MODAL
INPUT ON EFL ACADEMIC LISTENING COMPREHENSION



HÜMEYRA GENÇ

T.C. YEDITEPE UNIVERSITY

2017

The Impact of Context and Content Visuals in Multi-Modal Input on EFL Academic
Listening Comprehension

Dissertation Submitted to the Institute of Educational Sciences in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy
in
English Language Education

by
Hümeýra Genç

Yeditepe University



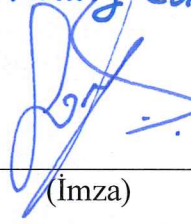
YEDİTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ

TEZ TESLİM ve ONAY TUTANAĞI


KONU: *The Impact of Context and Content Visuals in Multi-modal Input on EFL Academic Listening Comprehension*

ONAY:


Yard. Doç. Dr. Gonca Cumaoğlu
(Danışman)


(İmza)

Prof. Dr. Arif Altun
(Üye)


(İmza)

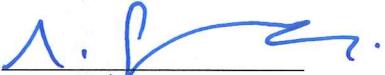
Doç. Dr. Hossein Farhady
(Üye)


(İmza)

Yrd. Doç. Dr. Zeynep Koçoğlu
(Üye)


(İmza)

Yrd. Doç. Dr. Adem Soruç
(Üye)


(İmza)

TESLİM EDEN

TEZ SAVUNMA TARİHİ

TEZ ONAY TARİHİ

: *Hümeyra GENÇ*
: *26.04.2017*
: *26.04.2017*

CURRICULUM VITAE

HÜMEYRA GENÇ

EDUCATION

- 2008 -2017** **PhD. Degree in English Language and Education**
The Institute of Educational Sciences, Department of English Language
Education, Yeditepe University, Istanbul, Turkey
- 2002-2004** **Post-Graduate Degree, Telecommunication for Multimedia and Internet-
based Instruction**
School of Professional Studies in Business and Education,
Johns Hopkins University, Baltimore, MD, USA
- 2000-2004** **Master of Science Degree in General Educational Studies**
School of Professional Studies in Business and Education,
Johns Hopkins University, Baltimore, MD, USA
- 1994-1998** **B.A. Degree in English Language Teaching,**
Department of Foreign Language Education,
Uludağ University, Bursa, Turkey

PAPERS PRESENTED IN INTERNATIONAL CONFERENCES

- Genç, H. (2014, May). *A Critical Overview of English Language Teacher Education in Turkish Education System: Preservice and In-service*. Paper was presented at the 4th International Conference on Foreign Language Teaching and Applied Linguistics (FLTAL). Sarajevo, Bosnia and Herzegovina.
- Genç, H. (2014, May). *Deconstruction and Reconstruction Attempts: The Curriculum of Turkish Ministry of National Education and 6th Grade English as a Foreign Language Program*. Paper was presented at the 4th International Conference on Foreign Language Teaching and Applied Linguistics (FLTAL). Sarajevo, Bosnia and Herzegovina.

- Genç, H. (2014, March). *Second Language Hypertext Reading: The Investigation of Strategy Use*. Paper was presented at the International Academic Conference on Education, Teaching, and E-learning (IAC-ETeL), Prague, Czech Republic.
- Genç, H. (2013, May). *A Comparative Study: Syntactic Awareness in Young Turkish Monolingual and Turkish-English Bilingual Children*. Paper was presented at the International Journal of Arts & Sciences (IJAS) conference, Boston, MA, USA.
- Genç, H. & Gulozer, K. (2012, June). *Cognitive Load Effects of Instructional Formats and Presentation Types on Students' L2 Reading Comprehension Performance*. Paper was presented at the EdMedia World Conference on Educational Media and Technology Conference, Denver, CO, USA.
- Genç, H. (2012, June). *An Evaluation Study of a CALL Application: With BELT or Without BELT*. Paper was presented at the EdMedia World Conference on Educational Media and Technology Conference, Denver, CO, USA.
- Genç, H. (2011, March). *Investigating In-Service EFL Teachers' Attitudes towards Integrating ICT into Instruction*. Paper was presented at 22nd International Society for Informational Technology and Teacher Education Conference, Nashville, TN, USA
- Genç, H. & Gulozer, K. (2011, March). *An Experimental Study: The EFL Pre-service Teachers' English Language Writing Experience through Wiki*. Paper was presented at 22nd International Society for Informational Technology and Teacher Education Conference, Nashville, TN, USA
- Genç, H. & Gulozer, K. (2010, May). *A Collaborative study: The Effects of Online Discussions and Feedbacks on English Language Writing Performance*. Paper was presented at The 6th ELT Research Conference Ephesus, Selçuk, Turkey.

- Genç, H. (2010, May). *Paper and Screen: Reading Strategies Used by Low-proficient EFL Learners*. Paper was presented at The 6th ELT Research Conference Ephesus, Selcuk, Turkey.

ARTICLES PUBLISHED IN REVIEWED JOURNALS AND PROCEEDINGS

- Genç, H. (2014, March). Second language hypertext reading: The investigation of strategy use. Proceedings of IAC-ETeL 2014, ISBN 978-80-905442-3-9.
- Genç, H. (2014). A Critical Overview of English Language Teacher Education in the Turkish Education System: Pre-service and In-service. In A. Akbarov, *The Practice of Foreign Language Teaching: Theories and Applications* (pp. 227-235). Newcastle, UK; Cambridge Scholars Publishing.
- Genç, H. & Gulozer, K. (2013a). The effect of cognitive load associated with instructional formats and types of presentation on second language reading comprehension performance. *TOJET: The Turkish Online Journal of Educational Technology- TOJET*, 12(4), 171-182.
- Genç, H. (2013b). Deconstruction and reconstruction attempts: The curriculum of Turkish ministry of national education and the sixth grade English as a foreign language program. *Sino-US English Teaching*, 10(8), 593-607.
- Genç, H. (2013c). A comparative study: syntactic awareness in young Turkish monolingual and Turkish-English bilingual children. *Journal of Teaching and Education*, 2(2), 413-422.
- Genç, H. (2012). An Evaluation Study of a CALL Application: With BELT or without BELT. *Turkish Online Journal of Educational Technology-TOJET*, 11(2), 44-54.
- Genç, H. (2011). Paper and Screen: Reading Strategies Used by Low-proficient EFL Learners. *Sino-US English Teaching*. 8(10), 648-658.

- Genç, H. (2011). Investigating In-service EFL Teachers' Attitudes towards Integrating ICTs into Instruction. In Proceedings of Society for Information Technology & Teacher Education International Conference 2011 (pp. 2465-2472). Chesapeake, VA: AACE. Retrieved from <http://www.editlib.org/p/36680>
- Genç, H. & Gulozer, K. (2011). An Experimental Study: The EFL Pre-service Teachers' English Language Writing Experience through Wiki. In Proceedings of Society for Information Technology & Teacher Education International Conference 2011 (pp. 1498-1503). Chesapeake, VA: AACE. Retrieved from <http://www.editlib.org/p/36507>



ACKNOWLEDGMENTS

Completing this dissertation has been a life changing experience for me and I believe that it would not have been possible to do this without the assistance of several people who have touched my life. Therefore, there are so many invaluable people to thank for their support, and contribution along my PhD journey.

First of all, I would like to express my deepest gratitude to my supervisor, Assist. Prof. Dr. Gonca KIZILKAYA CUMAOĞLU for her unwavering support, constructive feedback, and valuable advice over the past four years. Without her continuous guidance, I would not have managed to complete my dissertation.

I am also grateful to Prof. Dr. Ayşe Semra AKYEL for her academic guidance and constant encouragement. I learnt a lot from her throughout my PhD study.

I would like to express my sincere thanks to the members of my dissertation committee, Prof Dr. Arif ALTUN, and Assist. Prof. Dr. Zeynep Banu KOÇOĞLU. I am extremely grateful for their assistance, important comments, and valuable suggestions throughout my study.

I would also like to extend my appreciation to Assoc. Prof. Dr. Hossein FARHADY. The designing part of my study would not have been completed without the assistance of him.

I wish to express my gratitude to one of the members of my jury, Assist. Prof. Dr. Adem SORUÇ, for his useful suggestions.

I owe special thanks to Sabahattin ATALAY and Damla ERKUL for supporting me to conduct my study in the university where they were the directors of the English Language Institute and for designing their syllabi based on my research needs.

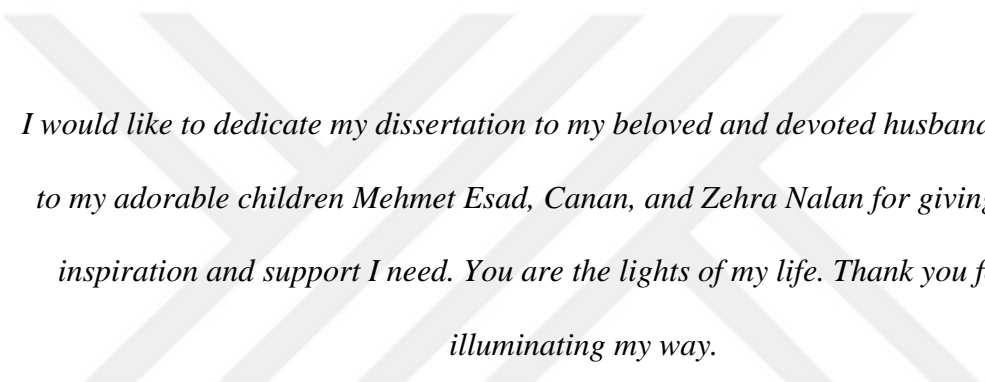
In particular, I would like to express my special thanks to those who have offered professional support, fruitful opinions, positive criticisms, and invaluable friendship over the years; Sevgi ŞANLITÜRK, Yeliz ÖZTOP, Neslihan GÜNAYDIN, and Magdalena ADAMCZYK. They never gave up believing in me and I am sure that much of the construction part of my data collection instruments would not have been accomplished without their precious help.

I would also like to express my gratitude to a couple of colleagues and my friends for helping and supporting me throughout this process: Süleyman ÇALIŞKAN for encouraging me to apply for a PhD program; Yurdanur TEKTAŞ for her warm heart, endless encouragement, and generous support; Ahmet Ramazan SALMAN for his joyful conversations; Onur Kadir URAL for his generous help in piloting the process of my data collection instruments; and all the instructors in the English language institute for allowing me into their classes to collect my data.

I am also grateful to my dearest friend Emine Esin EŞMEN whose emotional support and endless encouragement helped me cope with all the stress throughout my study and she never let me give up.

My sincere appreciation goes to my mother and father, Ayten and Kazım KARAKOÇ, for their continuous prayers, endless trust, and unconditional love. Your presence in my life is divine.

Finally, I am fully indebted to my husband, Ahmet Genç, and to my children, Mehmet Esad, Canan, and Zehra Nalan. You bless me with a life of joy in every moment when I am with you. You have always pushed me farther than I thought I could go. Without your understanding, patience, and ceaseless support I would not have completed this journey. I believe that I can overcome every difficulty with your love. I love you so much.



I would like to dedicate my dissertation to my beloved and devoted husband Ahmet, and to my adorable children Mehmet Esad, Canan, and Zehra Nalan for giving me all the inspiration and support I need. You are the lights of my life. Thank you for always illuminating my way.

TABLE OF CONTENTS

CURRICULUM VITAE	i
ACKNOWLEDGMENTS	v
DEDICATION	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xiv
LIST OF FIGURES	xvi
LIST OF ABBREVIATIONS	xvii
ABSTRACT	xviii
KISA ÖZET	xx
CHAPTER 1-INTRODUCTION	1
1.1. Background to the Study	1
1.2. Statement of the Problem	3
1.3. Purpose of the Study	5
1.4. Significance of the Study	7
1.5. Overview of the Methodology	8
1.6. Definitions of Significant Terms	10
1.7. Organization of the Study	12
CHAPTER 2-LITERATURE REVIEW	14
2.1. Listening	14
2.2. Listening Comprehension	15

2.3. Second and Foreign Language Listening Comprehension	17
2.3.1. Decoding and Meaning Building	22
2.3.2. Models of Listening Process	24
2.3.2.1. The Bottom-up Model	24
2.3.2.2. The Top down Model	25
2.3.2.3. The Interactive Model	27
2.4. Cognitive Theory of Multimedia Learning	29
2.5. Real-life Listening	31
2.6. Technology in Language Classroom	32
2.7. Academic Listening	34
2.8. Testing Listening	36
2.8.1. Background Knowledge	39
2.8.2. Text Characteristics	41
2.8.3. Speech Rate and Accent	42
2.8.4. Task Type	45
2.8.5. Language Proficiency Level	47
2.8.6. Visuals	49
2.8.7. The Length of Texts	50
2.8.8. Authentic Recordings	51
2.9. Role of Visuals in Second and Foreign Language Listening Comprehension.....	53
2.9.1. Types of Visuals	56
2.9.2. Types of Input Modes.....	57
2.10. Studies on the Impact of Visuals in Listening Comprehension.....	59
CHAPTER 3-METHODOLOGY.....	67
3.1. Research Questions	67

3.2. Research Design	68
3.3. Setting	71
3.4. Participants	73
3.5. Data Collection Instruments	77
3.5.1. Demographic Information and Information English Listening Questionnaire	77
3.5.2. EFL Listening Style Scale	77
3.5.2.1. Developing the EFL Listening Style Scale	78
3.5.2.2. Piloting and Validating the EFL Listening Style Scale	80
3.5.2.3. Reliability Study of the Scale	84
3.5.3. Listening Tests	85
3.5.3.1. Developing of the Listening Tests	85
3.5.3.2. Piloting the Listening Tests.....	103
3.5.3.3. Data Analysis of the Piloted Listening Tests.....	104
3.5.4. Interview.....	117
3.6. Data Collection Procedure.....	118
3.7. Data Analysis.....	122
CHAPTER 4- RESULTS.....	127
RQ.1. Is there any significant difference among different types of input modes, namely audio-only, audio with visuals and video, in the listening tests in terms of their impact on EFL students' academic listening performance?.....	127
RQ. 2.1. Is there any significant difference between audio with context visuals and audio with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?	130

RQ. 2.2. Is there any significant difference between video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?	131
RQ. 2.3. Is there any significant difference among audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?	132
RQ. 3.1. Is there any significant effect of proficiency level on EFL students' performance scores in academic listening subtests?.....	135
RQ. 3.2. Is there any significant effect of gender on EFL students' performance scores in academic listening subtests?	140
RQ. 3.3. Is there any significant effect of listening style on EFL students' performance scores in academic listening subtests?	141
RQ.4. What are EFL students' experiences about the presence and absence of context and content visuals provided by multi-modal input in the listening tests?.....	143
CHAPTER-5-DISCUSSION AND CONCLUSION	151
Research Question 1.....	151
Research Question 2.....	154
Research Question 3.....	162
Research Question 4.....	169
Conclusion.....	173

Implications.....	176
Limitations.....	179
Recommendations for Further Research.....	182
REFERENCES.....	184
APPENDICES.....	200
APPENDIX A Demographic Information and Information of English Listening	
Questionnaire (in English)	200
APPENDIX B Demographic Information and Information of English Listening	
Questionnaire (in Turkish)	202
APPENDIX C A Letter for Expert Opinion of EFL Listening Scale (in Turkish).....	204
APPENDIX D English as Foreign Language Listening Style Scale (in Turkish).....	208
APPENDIX E The Final Version of English as Foreign Language Listening Style	
Scale (in English).....	210
APPENDIX F The Final Version of English as Foreign Language Listening Style	
Scale (in Turkish)	212
APPENDIX G The Audio script of Part-1 (Context Audio) in Listening Test-1.....	214
APPENDIX H The Audio script of Part-2 (Content Audio) in Listening Test-1.....	216
APPENDIX I The Audio script of Part-3 (Context Video) in Listening Test-1.....	218
APPENDIX J The Audio script of Part-4 (Content Video) in Listening Test-1.....	220
APPENDIX K The Audio script of Part-5 (Audio-Only) in Listening Test-1.....	222
APPENDIX L The Audio script of Part-1 (Audio-Only) in Listening Test-2.....	224
APPENDIX M The Audio script of Part-2 (Content Audio) in Listening Test-2.....	226
APPENDIX N The Audio script of Part-3 (Context Video) in Listening Test-2.....	228
APPENDIX O The Audio script of Part-4 (Content Video) in Listening Test-2.....	230

APPENDIX P	The Audio script of Part-5 (Context Audio) in Listening Test-2.....	232
APPENDIX Q	The First Version of the Listening Test-1.....	234
APPENDIX R	The First Version of the Listening Test-2.....	245
APPENDIX S	The Final Version of the Listening Test-1.....	256
APPENDIX T	The Final Version of the Listening Test-2.....	267
APPENDIX U	Semi-Structured Group Interview Questions	278



LIST OF TABLES

Table 3.1 Data Sets Collected for the Study.....	69
Table 3.2 The Design of the Study.....	70
Table 3.3 The Design of the Study.....	70
Table 3.4 Background Information about the Participants (n=127).....	75-76
Table 3.5 Results of KMO and Bartlett's Tests.....	81
Table 3.6 Factor Analysis Results of the Scale-Rotated Components Matrix.....	83
Table 3.7 Reliability Values for the Scale.....	85
Table 3.8 Test Specifications of the Listening Tests.....	86
Table 3.9 Readability Statistics for the Scripts of Audios and Video Clips in the Tests.....	99
Table 3.10 Item Facility and Item Discrimination for Each Item on the Listening Test-1.....	105
Table 3.11 Item Facility and Item Discrimination for Each Item on the Listening Test-2.....	107
Table 3.12 Total Number of the Questions in each Subtest of both Listening Tests After Item Analysis.....	110
Table 3.13 Results of Distractor Analysis of Item Scores on the Listening Test-1.....	111
Table 3.14 Results of Distractor Analysis of Item Scores on the Listening Test-2.....	112
Table 3.15 Results of Reliability Analysis of the Overall Scores and Scores of each Part on the Listening Test-1 and the Listening Test-2.....	114
Table 3.16 Data Analyses Used for Answering the Research Questions.....	124
Table 4.1 Descriptive Statistics of Three Input Modes from the Listening Tests.....	128
Table 4.2 Repeated Measure ANOVA Test Result for Three Input Modes.....	129

Table 4.3 Bonferroni Test Results for Multiple Comparisons.....	130
Table 4.4 Paired-Samples t-Test Results for Audio with Context and Content Visuals.....	130
Table 4.5 Paired-Samples t-Test Results for Video with Context and Content Visuals.....	131
Table 4.6 Descriptive Statistics of Five Input Modes in the Listening Tests.....	132
Table 4.7 Repeated Measure ANOVA Test Result for Five Input Modes.....	133
Table 4.8 Bonferroni Test Results for Multiple Comparisons.....	135
Table 4.9 Descriptive Statistics of Three Proficiency Levels According to the Subtests.....	137
Table 4.10 A one-way ANOVA Result Table for the Subtests According to Proficiency Level.....	138
Table 4.11 Tukey Post-hoc Test Results for Multiple Comparisons According to Proficiency Level.....	139
Table 4.12 Independent Samples T-Test Results for the Subtests According to Gender.....	140
Table 4.13 Descriptive Statistics of Listening Styles According to the Subtests.....	142
Table 4.14 A one-way ANOVA Result Table for the Subtests According to the Listening Styles.....	143
Table 4.15 Participants' Positive Opinion about the Visuals in the Listening Tests...	144
Table 4.16 Participants' Negative Opinions about the Visuals in the Listening Tests..	145
Table 4.17 The Participant' Opinions about the Assistance of the Type of the Visuals in the Listening Tests (Total f= 24).....	146
Table 4.18 The Participant' Preference of the Type of the Visuals in a Listening Test. (Total f= 24).....	148
Table 4.19 The Participant' Preference of the Type of the Input Modes in a Listening Test. (Total f= 24).....	149

LIST OF FIGURES

Figure 1.1. Processing of aural and visual information in Cognitive Theory of Multimedia Learning.....	31
Figure 1. 2. A product approach to using questions in listening.....	37
Figure 1.3 A process approach to using questions in listening	37
Figure 3. 1. Number of the participants in the study in 2nd and 3rd module.....	74
Figure 3.2. Eigenvalue graphic.....	84
Figure 3.3. Characteristics of the parts in the listening tests.....	88
Figure 3.4. The input modes in each listening test.....	89
Figure 3.5. A screenshot from audio with content visuals in Test-1.....	94
Figure 3.6. A screenshot from audio with content visuals in Test-2.....	94
Figure 3.7. A screenshot from audio with context visuals in Test-1.....	95
Figure 3.8. A screenshot from audio with context visuals in Test-2.....	95
Figure 3.9. A screenshot from video with context visuals in Test-1.....	96
Figure 3.10. A screenshot from video with context visuals in Test-2.....	96
Figure 3.11. A screenshot from video with content visuals in Test-1.....	97
Figure 3.12. A screenshot from video with content visuals in Test-2.....	97
Figure 3.13. The order of the subtests in listening test-1.....	100
Figure 3.14. The order of the subtests in listening test-2.....	100
Figure 3.15. Data collection procedure.....	122
Figure 4.1 Estimated marginal means of three input modes in the listening test-1 and the listening test-2.....	128
Figure 4.2. Estimated marginal means of fine input modes in the listening test-1 and the listening test-2.....	133

LIST OF ABBREVIATIONS

L2: Second Language

FL: Foreign Language

ESL: English as a Second Language

EFL: English as a Foreign Language

EFL-LSS: English as a Foreign Language Listening Style Scale

CBT: Computer-based Testing

TOEFL: Test of English as a Foreign Language

IELTS: International English Language Testing System

CTML: Cognitive Theory of Multimedia Learning

CALL: Computer Assisted Language Learning

ABSTRACT

The Impact of Context and Content Visuals in Multi-modal Input on EFL Academic
Listening Comprehension

by

Hümeyra Genç

The purpose of this dissertation study is to investigate the impact of context and content visuals in multi-modal inputs on English as a foreign language students' academic listening comprehension. The study specifically aimed to examine; a) whether there exists a difference in EFL students' listening performance on five listening subtests, namely, audio-only mode, audio with context visuals mode, audio with content visuals mode, video with context visuals mode, and video with content visuals mode, b) whether individual differences, that is English language proficiency level, gender, and students' listening style, affect EFL students' academic listening comprehension. In doing so, various data were collected through four types of data collection instruments. First, EFL students' personal information was collected by a demographic information questionnaire, second, their listening styles were identified by EFL Listening Style Scale, then, their listening test scores on five different input modes (audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals) were collected by two listening tests, and finally, their opinions about the listening tests gathered by semi-structured group interview were determined. One hundred-twenty-seven EFL students from an English Language Institute of a private university in Istanbul, Turkey participated in this study.

The findings of the present study demonstrated that the inclusion of nonverbal information, whether context or content, via audio and video input modes in an

academic listening comprehension test has a positive effect, compared to audio-only mode, on EFL students' academic listening comprehension. Moreover, although the EFL students' English language proficiency levels have different effects on their listening performance to varying degrees, their gender and listening styles do not have any effect on their academic listening performance. The findings of the study provide significant practical and pedagogical implications for EFL teachers, ELT publishers, and test designers with respect to English language listening skill instruction, designing English language teaching materials, and developing English language listening tests.

Key Words: *Listening comprehension, context visuals, content visuals, video-mediated listening, multi-modal input, academic listening, listening style, testing listening.*

KISA ÖZET

Çoklu Biçimli Girdilerde Bağlam ve İçerik Görsellerinin Yabancı Dil Olarak İngilizce Akademik Dinleme Anlama Üzerindeki Etkisi

Hümevra Genç

Bu tez çalışmasının amacı, çoklu biçimli girdilerde bağlam ve içerik görsellerinin İngilizce yabancı dil öğrencilerinin akademik dinleme anlayışı üzerindeki etkisini araştırmaktır. Çalışma, özellikle şunları incelemeyi amaçlamaktadır; a) beş dinleme alt testinde- ki bunlar yalnız ses modlu, bağlamsal görselli ses modlu, içerik görselli ses modlu, bağlamsal görselli görüntü modlu ve içerik görselli görüntü modlu dinleme testleridir- İngilizce yabancı dil öğrencilerinin dinleme performanslarında bir farklılık olup olmadığı. b) bireysel farklılıkların, yani İngilizce dil yeterlilik seviyesinin, cinsiyetin ve öğrencilerinin dinleme stillerinin İngilizce yabancı dil öğrencilerin akademik dinleme anlayışlarını etkileyip etkilemediği. Bunu yaparken, pek çok veri dört çeşit veri toplama aracı yardımıyla toplanmıştır. İlk olarak, İngilizce yabancı dil öğrencilerinin kişisel bilgileri demografik bilgi anketi ile toplanmıştır. İkinci olarak, öğrencilerin dinleme stilleri İngilizce Yabancı Dil Dinleme Stili Ölçeği ile belirlenmişti. Sonra öğrencilerin beş farklı giriş modundaki (yalnız sesli mod, bağlamsal görselli sesli mod, içerik görselli sesli mod, bağlamsal görselli görüntülü mod ve içerik görselli görüntülü mod) dinleme testi puanları iki farklı dinleme testi ile toplanmıştır ve son olarak öğrencilerin yarı yapılandırılmış grup görüşmesi aracılığıyla dinleme testleri hakkındaki görüşleri belirlenmiştir. Bu araştırmaya, İstanbul'da özel bir üniversitenin İngilizce Dil Enstitüsü'nden gelen 127 İngilizce yabancı dil öğrencisi katılmıştır.

Bu çalışmanın bulguları, ses ve görüntü girdi modları yoluyla, bağlam ya da içerik olsun, sözel olmayan bilgilerin akademik bir dinleme anlama testine dâhil edilmesinin, İngilizce yabancı dil öğrencilerinin akademik dinleme anlayışında, salt ses moduna kıyasla, olumlu bir etkisinin olduğunu göstermiştir. Ayrıca, İngilizce yabancı dil öğrencilerinin İngilizce dil yeterlilik düzeyleri dinleme performansları üzerinde farklı derecelerde farklı etkilere sahip olsa da, cinsiyetleri ve dinleme stillerinin akademik dinleme performansları üzerinde herhangi bir etkisi yoktur. Çalışmanın bulguları, İngilizce dinleme becerisi öğretimi, İngilizce öğretim materyalleri tasarlanması ve İngilizce dil dinleme testleri geliştirilmesi alanlarında İngilizce yabancı dil öğretmenleri, İngilizce dil öğretimi materyal tasarımcıları ve yayıncıları ve test tasarımcıları için önemli pedagojik ve uygulanabilir sonuçlar doğurmaktadır.

Anahtar Kelimeler: *Dinlediğini anlama, bağlamsal görseller, içerik görselleri, video aracılı dinleme, çoklu biçimli girdi, akademik dinleme, dinleme stili, dinlemenin test edilmesi.*

CHAPTER 1

INTRODUCTION

1.1. Background to the Study

In real life situations, listeners are generally faced with the physical presence of the speakers they are listening to and the verbal information is generally accompanied by visual information (Suvorov, 2013; Wagner, 2010b). In real life listening, the visibility of the speaker and the context in which the speech takes place are inevitable unless we are on the phone or listening to the radio. People are so surrounded by the environmental or visual clues (e.g., facial expressions, gestures, mimics, tone of voice, pictures, charts, maps, diagrams, stills) when they listen to something that most of the time they are not aware of how much these clues help them comprehend the spoken messages. According to Ur (2009), although noises and smells also support background information to comprehend the incoming message, most of the clues around us are visuals. She also points out that sound recordings, radio broadcasts, and telephone talks have little clues that contribute to comprehension in our daily lives and these kind of listening situations comprise only a small part of our total listening activity. Therefore, people generally find the listening without visual supports “cold” and when we listen to something which does not have any visual clues, it takes us sometime to “tune in” our ears to the content of what we hear (Vandergrift and Goh, 2012).

Demand for learning English is growing very fast because of the rapid pace of globalization. Individuals who are looking for university education or any employment positions especially in English speaking countries have to attend English language schools in order to improve their level of English. With the advancement in

multimedia technologies in computer technology, language schools integrate more realistic aids into their instruction, especially into listening lessons. The inclusion of visual elements in second language (L2) and foreign language (FL) listening instruction improves the authenticity of classroom practice because in real-life situation we are usually surrounded by different types of visuals. Ur (2009, p. 5) states that “if the speaker is usually present in real life listening situations, ...perhaps we should think again about how much we ought to use recordings as the basis of our exercises. Perhaps we should revert to using live speakers...”. Moreover, Vandergrift and Goh (2012) mention that L2 language learners at lower-level of language proficiency benefit from listening activities which are supported by visual images in language instruction. Similarly, advanced level language learners can manage visuals and audio input together better even though there are sometimes mismatches between them.

Regarding the inclusion of real life materials in L2 listening instruction, publishing companies and software companies have been putting a great deal of effort into developing materials that include especially videos and animations for the last decade because including them in the instruction provides face to face contact in which the listener can see the speaker.

Moreover, most of the academic skill coursebooks, not only for listening but also for reading and speaking skills, comprise visual supports like photos, pictures, maps, drawings, videos, or charts in order to support L2 language learning in a more realistic way. According to numerous L2 and FL language instructors, the inclusion of visuals to audio texts is considered a better representation of actual communicative situations in classroom environment. Furthermore, a few international standardized English language tests such as Test of English as a Foreign Language (TOEFL) or

International English Language Testing System (IELTS) have started to include visual supports to provide more reliable and valid tests as well. As a result, although the inclusion of real-life elements like visual inputs in language learning materials is becoming increasingly widespread, they are still integrated into language tests cautiously. Therefore, some researchers (e.g. Coniam, 2001; Ginther, 2002; Gruba, 1993; Suvorov,2009; Wagner, 2008; Wagner, 2010b) have been investigating the effects of visuals on language learning and on language testing in recent years.

1.2.Statement of the Problem

Although the inclusion of visuals in L2 language learning (especially in L2 listening comprehension) seems very important (Rost,2011), they are rarely used in assessing students' L2 listening ability. One of the most important reason why test developers do not include visuals frequently in L2 listening tests is the lack of agreement about what is assessed in L2 listening tests: the listening ability itself or the ability to understand both visual and verbal information together.

Some researchers (Ockey, 2007; Wagner, 2007) believe that since the verbal interaction occurs with visual information in real life settings, L2 listening tests should involve visual information as well. However, the researchers who oppose this view believe that visuals must not be included in L2 listening tests because in the tests, the listening ability is tested not the ability to understand visual information (Buck, 2001). According to Suvorov (2013), these opponents consider that if they include the ability of using visual information in L2 listening tests, they will actually test more than the listening ability. However, the current developments in construct definitions in integrated language assessment propose that constructs should involve a range of different abilities (Cumming, 2013).

Another important reason for not including visual information in L2 listening tests is the inconclusive results of the effects of visuals on test-takers' performance on L2 listening tests. Some studies indicate that test takers perform better on tests which include visual input (Ginther, 2002; Wagner, 2010b) while others have found detrimental effect (Suvorov, 2009) or no effect (Coniam, 2001; Gruba, 1993; Ockey, 2007; Suvorov, 2013) on facilitating the understanding of verbal information on the performance of test takers on these test formats. The effects of visual input on the comprehension of verbal input in these research studies may depend on several factors such as text types, the types of the test tasks, the kinds of visual materials, the characteristics of the test-takers, or the interaction of all these factors.

However, the main reason of these mixed results concerning the impact of visual information on students' L2 listening comprehension can be associated with using different types of visuals in L2 listening tests. The types of visuals are separated as context (or Bejar, Douglas, Jamieson, Nissan, & Turner, 2000 defined context visuals as situational visuals) and content visuals in literature (Ginther, 2002). Context visuals or situational visuals involve visual information about the setting in which the verbal exchange takes place (e.g. a picture of an instructor giving a lecture in the classroom or a picture of the classroom where the lecture is given). Contrarily, content visuals comprise visual information that is related to the content of the verbal interaction (e.g. a picture of a diagram that shows the effects of climate change in the world). The result of reviewing research studies in the literature shows that the visual type that is most frequently used is context visuals. Content visuals are rarely implemented in research studies. Ginther, (2002) and Suvorov (2013) are the only researchers who investigated the difference between the context and content visuals in their studies. Since the impacts of visuals on L2 students' test performance can be attributed to the

different types of visuals that are involved in listening tests, further research is needed to determine the effects of different types of visuals, context and content, on students' performance on L2 listening comprehension test.

1.3. Purpose of the Study

Taking into account the inconclusive results of the current research studies comparing the effects of different types of visuals (context and content) on L2 students' listening test performance; the lack of investigation that analyzes the effects of these two types of visuals (context and content) delivered via different input modes (audio and video) on L2 students' academic listening comprehension; the lack of investigation of the effect of individual differences (gender, proficiency level, and listening style) on two types of visuals (context and content) on L2 students' academic listening comprehension, the aim of the present study is to investigate the impact of context and content visuals in multi-modal inputs (video, audio with visuals and audio-only) on English as a foreign language students' academic listening comprehension. Regarding this purpose, the present study aimed to address the following research questions:

1) Is there any significant difference among different types of input modes, namely audio-only, audio with visuals, and video, in the listening tests in terms of their impact on EFL students' academic listening performance?

2) Is there any significant difference between different types of visuals, context and content, in the listening tests in terms of their impact on EFL students' academic listening performance?

2.1) Is there any significant difference between audio with context visuals and audio with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

2.2) Is there any significant difference between video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

2.3) Is there any significant difference among audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

3) Is there any significant effect of individual differences (proficiency level, gender and listening style) on EFL students' performance scores in academic listening subtests, namely audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals?

3.1) Is there any significant effect of proficiency level on EFL students' performance scores in academic listening subtests?

3.2) Is there any significant effect of gender on EFL students' performance scores in academic listening subtests?

3.3) Is there any significant effect of listening style on EFL students' performance scores in academic listening subtests?

4) What are EFL students' experiences about the presence and absence of context and content visuals provided by multi-modal input in the listening tests?

1.4. Significance of the Study

First of all, by presenting the effects of context and content visuals in multi-modal input (audio with content visuals, audio with context visuals, video with content visuals and video with context visuals) on EFL students' academic listening comprehension, this study makes an important contribution to the field of foreign and second language teaching, FL and L2 assessment, FL and L2 material design, and computer-based testing to explain which type of visuals -context or content- affect EFL students' academic listening comprehension and, if they affect, how the inclusion of these types of visuals affect EFL students' academic listening comprehension on such listening tests.

Secondly, by comparing the impacts of different types of input modes, audio-only, audio with visuals, and video, on EFL students' academic listening comprehension, the present study makes a significant contribution to the field of FL and L2 assessment and computer-based testing to provide how different types of input modes influence EFL students' academic listening comprehension on listening tests.

Thirdly, examining the effects of individual differences (i.e., proficiency level, gender, and listening style) on EFL students' performance scores in academic listening subtests, which involves audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals, this study provides a significant contribution to the field of FL and L2 teaching and FL and L2 testing to clarify which individual difference affects EFL students' academic listening comprehension on listening tests and, if the comprehension is affected, how EFL students' academic listening comprehension on these listening tests is affected by these differences.

Furthermore, since there has been no attempt to elicit EFL students' English language listening style while listening in English language and to develop a scale, which aims to investigate EFL students' English language listening style, in the field of FL and L2 language learning, teaching, and testing so far, this study makes a great contribution to the field by developing an EFL Listening Style Scale.

Finally, the present study makes a significant contribution to the field of FL and L2 assessment, FL and L2 language learning and teaching, L2 material design, and computer-based testing by providing practical and pedagogical implications for L2 listening tests.

1.5. Overview of the Methodology

Participants

One hundred-twenty-seven EFL students from randomly selected classes, sixty-nine of them were female and fifty-eight of them were male, participated in the present study. In order to eliminate the effect of culture on listening comprehension, only Turkish students were involved in the study. At the time of the data collection, 42 students were in A2 level (elementary language level of English), 43 of them were in B1 level (intermediate language level of English), and 42 of them were in B2 level (upper-intermediate language level of English).

Setting

The study was conducted at the English Language Institute of a private university in Istanbul, Turkey in 2014-2015 academic years. This university was a thematic university focusing on social science with approximately 1,600 students. Besides, it had very limited number of departments with two faculties (faculty of management

and administrative science and faculty of humanities and social science), and two institutes (institute of social science and English language institute).

Data Collection Instruments

A Demographic Information Questionnaire, an EFL Listening Style Scale, two listening tests and a semi-structured group interview questionnaire were developed and used as data collection instruments for the present study. The EFL Listening Style Scale was piloted twice (n=300 and n=275) in order to validate the reliability and validity of the scale. Furthermore, the pilot study for the listening tests conducted with 61 EFL students in order to validate the reliability and validity of them.

Data Analysis

All statistical analyses of data collection instruments were performed by SPSS 20.0 and Microsoft Excel (2016). Frequencies, means and standard deviations were used for descriptive statistics. Item analysis including item facility and item discrimination and distractor analysis, were conducted in order to construct the best items and distractors for the listening tests. Factor analysis was used for the validity study of the EFL Listening Style Scale. Besides, reliability analyses were carried out in order to measure the reliability scores of the EFL Listening Style Scale and the Listening Tests1 and Listening Test-2.

With regard to the first research question which is “Is there any significant difference among different types of input modes, namely audio-only, audio with visuals and video, in the listening tests in terms of their impact on EFL students’ academic listening performance?” one way repeated measures-ANOVA was carried out.

Concerning the second research question, “Is there any significant difference between different types of visuals, context and content, in the listening tests in terms of their impact on EFL students’ academic listening performance?” a paired-samples *t* tests and one way repeated measures-ANOVA were conducted.

For answering the third research question, “Is there any significant effect of individual differences (proficiency level, gender, and listening style) on EFL students’ performance scores in academic listening subtests, namely audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals?” a one-way ANOVA and independent sample *t*-test were performed.

Finally, in order to analyze the qualitative data, which was aimed to answer the fourth research question, “What are EFL students’ experience about the presence and absence of context and content visuals provided by multi-modal in the listening tests?” content analysis technique was carried out. Qualitative interview data were translated into English and transcribed for coding manually.

1.6. Definitions of Significant Terms

Context Visuals: are visuals that provide information about the context for the verbal exchanges, such as the images of the participants and/or the setting (Ginther, 2002).

Content Visuals: are visuals that are related to the content of the verbal interaction and may include still photos, pictures, drawings, or diagrams (Ginther, 2002).

Multi-modal Input: is a recording of a spoken text that includes both the audio and visual input together. Video is an example of this kind of input.

Input Modes: are types of channels which are used to deliver an information to a listener (e.g., audio / video)

Modes: refer to the method of presentation in which information is represented and shared, such as through audio, video, or text (Mayer, 1997).

Still Image: is a single static image. It is non-moving visual information such as pictures or graphs.

Audio-Only: the text is presented to the listeners solely through the aural channel (Wagner, 2013).

Audio with Visual: both the aural and visual channels with still images are used together.

Video: includes a prerecorded motion pictures that involve moving and dynamic images together with auditory input.

Background (Prior) Knowledge: information that is essential to understanding a situation or a problem.

Academic Language: is a set of thinking skills and abilities to encode and decode complex thoughts (Diaz-Rico and Weed, 2002 cited in Zwiers, 2007).

Listening Comprehension: is a complex communication behavior, involving process of receiving, attending to and assigning meaning to verbal and/or non-verbal (visual) stimuli (Coackley and Wolvin, 1986).

Academic Listening Comprehension: is a processing of discipline-specific information which students get from visual and auditory clues in order to define what is going on and what the speakers are trying to express in an academic context.

Learning Style: is the internally based characteristics of individuals for the intake or understanding of new information. (Reid 1995).

Listening Style: is the fairly fixed and internally based characteristics of listeners for understanding incoming aural message.

Visual Style: is a fixed characteristic of a L2 listener who understands L2 verbal information with the help of visual information best. This listener needs the support of visuals in order to make sense of the L2 verbal input.

Auditory Style: is the fixed characteristic of a L2 listener who understands L2 verbal information only through auditory input. This listener does not need any aid other than auditory input to understand the verbal message.

Spatial Style: is the fixed characteristic of a L2 listener who understands L2 verbal information through visualizing the mental representation of the content of the verbal message in her mind.

Bottom-up Style: is the fixed characteristic of a L2 listener who understands L2 verbal information through lower level decoding process until the meaning is built.

1.7. Organization of the Study

This dissertation consists of five chapters; an introduction, a literature review, methodology, the results, and a discussion and conclusion. In the first chapter, a general introduction to the study is provided by presenting brief background information, statement of the problem, purpose of the study, research questions, significance of the study, overview of methodology (participants, setting, data collection instruments, and data analysis), and definitions of significant terms.

The second chapter presents a detailed literature review organized under ten main headings; listening, listening comprehension, second and foreign listening comprehension, cognitive theory of multimedia learning, real-life listening,

technology in the language classroom, academic listening, testing listening, role of the visuals in second and foreign language listening comprehension, and studies on the impact of visuals in listening comprehension.

In the third chapter, research questions, research design of the study, setting, participants, data collection instruments, data collection procedure, and data analysis were presented.

In the chapter four, the results of the research questions were presented. Finally, chapter five discusses and concludes the findings of the study along with implications, limitations and suggestions for further research., references and appendices were also included at the end of the dissertation.

CHAPTER 2

LITERATURE REVIEW

In order to provide a conceptual framework for the dissertation, this chapter introduces a literature review which involves subchapters about the definition of listening, L1, L2, and FL listening comprehension, models of listening comprehension, cognitive theory of multimedia learning, real life listening, testing listening, academic listening, factors that affect listening comprehension in testing, role of visuals in L2 listening comprehension, and studies on impact of visuals in listening comprehension.

2.1. Listening

There exist several different definitions of listening which have been evolving in time in the literature. Among various definitions, Lado (1961) provides the earlier definition of listening as “recognition control of the signaling elements of the language in communication situations” (p. 206). Besides, as a simplified definition, Underwood (1989) defines listening comprehension as "the activity of paying attention to and trying to get meaning from something we hear" (p. 1). Brown (2011), in his book, also gives a description of listening as “making sense of aural input” (p.5). In addition to these definitions on listening, several researchers include the role of visual information in the definition of it. For example, Coakley and Wolvin (1986) mention that listening is “a complex communication behavior, involving a process of receiving, attending to, and assigning meaning to verbal and/or non-verbal stimuli” (p. 20). Moreover, Purdy (1997) states that listening is "the active and dynamic process of attending, perceiving, interpreting, remembering, and responding to the expressed (verbal and nonverbal), needs, concerns, and information offered by other human

beings" (p. 8). Furthermore, Rubin (1995) defines listening as "consists of processing information which listeners get from visual and auditory clues in order to define what is going on and what the speakers are trying to express" (p. 151). Regarding to the definitions of listening, although the earlier definition of it (e.g., Lado, 1961) focuses on the role of aural input, the later definitions include the role of visual input in the definitions. Several researchers claim that listening comprehension is not limited to inclusion of aural input, but the function of visual input in listening process is also very important (Gruba, 1997; Ockey, 2007; Rubin, 1995).

2.2. Listening Comprehension

Although, listening was seen as a passive process, contrary to common assumption, now it is admitted as a dynamic process (Vandergrift, 2004) in which listeners are very active to comprehend a spoken message while listening (Vandergrift and Goh, 2012). Furthermore, although it is a widely accepted view in the research that conducting an effective listening research is difficult and complex issue because of the inaccessibility of what happens in listener's mind (Lynch, 2009) and also listening, which involves a meaning making process, is found challenging because of not seen directly (Brown, 2011). The author also explains that making sense refers to something that takes effort. In order to make sense what we hear, we have to use knowledge of the language such as sounds, words and grammar, topic knowledge, context knowledge and background knowledge (Brown, 2011).

Parallel to Brown, Vandergrift and Goh, (2012) mention that a listener uses different types of knowledge including linguistic knowledge, pragmatic knowledge, prior knowledge, and discourse knowledge in order to comprehend spoken message. The authors express that linguistic knowledge involves phonological and syntactic knowledge which are the main elements for listening comprehension. The authors

also indicate that pragmatic knowledge is very important to understand beyond the literal meaning of words and the intended meaning of the speakers. Another type of knowledge mentioned in the book of Vandergrift and Goh is that prior knowledge. When verbal information (the linguistic input) is heard by a listener, she always tries match this input with what she knows about it (their prior knowledge) in order to make the input sense. The last knowledge stated in the book is discourse knowledge. This knowledge help listeners know how different text types are created or how information is organized in certain texts. Vandergrift and Goh (2012) explain how listeners' cognitive process work while listening as "they draw on their mental lexicon for the linguistic knowledge necessary to parse the input, and on their bank of prior, pragmatic, and discourse knowledge to interpret the overall intended meaning on their interlocutor within the context of the interaction" (p.28). Consequently, listeners use these knowledge sources in order to reach meaningful interpretation while listening (Vandergrift and Goh, 2012).

Compared to listening to a foreign language, listening to our native language needs less effort. According to Field (2008, p.2) "as infants, we acquire listening skills without being conscious of any cognitive demands being made upon us." Besides, most of the human beings start to hear sounds even in their mothers' womb (Flowerdew and Miller, 2013). In the first language, knowledge that we need in order to comprehend the aural input is implicit because we acquire them automatically as we grow up and we all know the rules of our native language and use them correctly (Buck, 2001). On the other hand, L2 language learners need more conscious effort when they listen but if they gain more language experience and improve their language proficiency, they can process information fast and unconsciously, and also reach knowledge sources more quickly (Vandergrift and Goh, 2012).

2.3. Second and Foreign Language Listening Comprehension

Until 1970s, listening research had focused on researching listeners' first language not their second language because researchers believed that listening in second language was the same as in the listeners' first language. Therefore, what had been done in researching first language listening was helpful for understanding second language listening (Flowerdew and Miller, 2010). Vandergrift and Goh (2012) state that all L2 learners already acquire listening ability, which they need in learning L2, in their first language. The authors also mention that this ability that we get through L1 acquisition contribute to L2 listening ability. Flowerdew and Miller (2013) mention that comprehension of the spoken information is the only goal in L1 listening. Listener remakes the spoken message, checks the speakers' intention and extracts meaning from the message (Field, 2008). However, in L2 listening comprehension, listeners have two goals. The first one is to understand the aural message and the second one is to learn the target language and increase their language proficiency through listening activities. In similar line, Brown (2011) mentions that although children learn first speaking in L1 situation, they learn speaking commonly after reading in L2 context, especially in EFL situation. Therefore, the author finds difficult to explain the connection between oral and written abilities through L2 research.

According to Field (2008), learners of L2 find listening skills as frustrating because their progress is generally slow, spoken language is intangible and is difficult to control compare to written language. Since L2 listeners have limited language knowledge in terms of language proficiency, they cannot process every spoken message automatically. However, in L1 listening, listeners automatically, successfully and very rapidly comprehend the message with little attention to individual words and

L1 listeners listen more consciously only if there are unknown words, noise, a different accent, or an unknown topic in the spoken message. Most of the L1 listeners meet with difficulties in listening when they learn a new language (Vandergrift and Goh, 2012). Hulstijn (2001) suggests listening to 'I minus 1 level' texts in contrast to Krashen's 'I plus 1' principle (Krashen, 1985). It is claimed that listening to 'I minus 1' text will be very easy for L2 listeners and they will be able to understand almost everything with little effort in the listening text, and this application will increase L2 learners' motivation towards listening skill (as cited in Vandergrift and Goh, 2012).

Parallel to Field (2008) and Flowerdew and Miller (2013), Ur (2009) points out that language listeners think that successful comprehension means understanding completely everything what they hear even the incoming message is totally unimportant. The author also continuous that especially in the early level of language learning, listeners are absolutely discouraged and lose their motivation when they face with unknown words. Whereas listening texts are getting longer, much of the content is redundant. Therefore, language listeners must accept that effective listening means skipping unimportant words and understanding the main message (Ur, 2009).

In his book, Field (2008) clarifies that reading and listening skills share common meaning building elements, and both draw upon the same comprehension process. Therefore, methodologies can be transferred from one skill to another, such as employing similar exercise types, in order to check learners' understanding. However, Lynch (2009) states that most of the people find listening a foreign language more difficult than reading it. Furthermore, despite having similarities, many scholars mention that there are marked differences in reading and listening skills (Brown, 2011; Field, 2008; Lynch, 2011; Rost, 1990). Especially the decoding process is completely different because listening involves the process of sounds but

reading requires the process of print (Lynch, 2011). Field (2008) mentions the differences between listening and reading skills in his book briefly as follows: (1) speech involves different sounds which vary from one speaker to another, (2) speech is transient in time, listener cannot go back and check her understanding, not recursive like reading, (3) since it is temporary, listener has to carry all expressed ideas in her memory, (4) in a speech, speaker may mispronounce the words, speak very fast, swallow her words, hesitate, rephrase, repeat and sometimes forget what she is going to say, (5) listening is internalized activity, it cannot be observable, and cannot be studied directly. In addition to Field (2008), Flowerdew and Miller (2013) mention that we can see more lexical density in written works, which means that the ratio of content words to grammatical words. Moreover, they add that spoken messages involve "... much more emotional nuance, contextual sensitivity, personal weighting, interactive hooks, and reference to the real world outside the text." (p.51).

The teaching of L2 listening receives little attention compared to teaching reading, writing, even speaking. Although teachers teach how to read or write or making oral presentation, they rarely teach how to listen in language classrooms (Rost, 1990 and Vandergrift and Goh, 2012). Field (2008), in his book, emphasizes that listening was used as a means of teaching grammar through model dialogues until late 1970s and listening parts placed in the Cambridge First Certificate exam from 1970. As Flowerdew and Miller (2013) state that nowadays different kinds of approaches are used to teach listening skills in our language classrooms. Moreover, various choices of textbooks which aim at specifically teaching listening skills are available on the market.

In today's world, access to listening resources is very easy compare to past. Since CDs and DVDs are produced a lot and are inexpensive, foreign language

learners can buy them easily for extensive listening practice. Besides, free downloadable mp3-4 files, audio books and podcasts are available on many web sites on the Internet. Although there is numerous choice of listening materials accessible, such as CDs, DVDs, multimedia programs, and downloadable mp3 or 4 files from the Internet in our language classrooms today, listening skill courses are still underestimated and have been little discussed or researched (Field, 2008). One of the reasons for this lack of priority is the difficulty of teaching this skill because compared to speaking and writing skills, listening is still seen as 'passive' and largely 'hidden' process (Brown, 2011). It occurs in listener's mind and is difficult to show demonstrable results and observable products even though she demonstrates some signs of responds that shows comprehension. That's why teachers did not begin to practice listening as a skill until late 1960s (Field, 2008; Lynch, 2009; Nunan, 1999) and they could see how active the process of listening is through research in the 1980s (Brown, 2011).

As stated by Field (2008), language learners feel themselves most insecure on listening because: there is no tangible evidence that students are making progress in acquiring the skill; listening takes place in real time, it is very difficult to keep the speech in memory if it is not understood at the moment; making mistakes especially at a lower level cause a loss of confidence. As a justification for this opinion, Ur (2009) points out that a foreign language learning is a tiring process, especially listening skill, because keeping all words, phrases and syntax in the memory and trying to interpret them accurately exhaust language learners. Contrary to other skills, reading, writing, and speaking, listeners cannot set their own pace because the pace is set by a speaker in the listening. Ur (2009) also adds that at the beginning of the long listening

texts, listeners' amount of comprehension is higher than the comprehension towards the end because listeners run out of energy.

Field (2008), in his book, states that there is a belief that a lot of exposure to second language listening improves students' listening comprehension. He explains that this situation may be true if language learners live in a target language environment because learners live in the environment where the language is spoken exposes a lot of language input. Thus, living in a such condition increases learners' motivation and they can easily extract meaning from the speech around them.

However, the result for learners who expose to a target language only in a classroom environment is mostly different. Although these learners meet with a series of recordings with comprehension questions in the classroom, they do not use the target language outside world. Therefore, some learners improve their listening skills but many other do not. Parallel to Field (2008), Ur (2009) mentions that the listening activities in language classrooms must have some features of real life listening. She adds that, however, the listening texts which are used in language classes today are nearly composed as a written text and recorded for in classroom use. Language learners listen these texts without knowing what they are going to listen and having a clear purpose of it and then they commonly try to answer comprehension questions. According to the author, this process is very favorable as classroom techniques because it gives listening practices to the language learners but it does not make students ready for real-life listening. This kind of imitation of reality is not enough for learners.

Parallel to Field (2008), Flowerdew and Miller (2013) state that although both first language and second language listening share similar comprehension process for learners, some barriers to second language listening such as outside noise or cultural

background knowledge make it a difficult task. Brown and Yule (1983) also express the factors affecting listening as; (1) speaker factor (number of the speakers, their accents, their rate of speech), (2) listeners factor (proficiency level of them, their interest to the topic), (3) the content (required background knowledge, the difficulty of grammar and vocabulary), (4) support (any visuals such as pictures, diagrams, or maps) (as cited in Nunan, 1999). Likewise, Lynch (2009) points out some additional factors that inhibit listening comprehension as; emotional breakdown, lack of interest, reaction to the speaker, preparing response to the speaker.

2.3.1. Decoding and Meaning Building

Both L1 and L2 listeners have to deal with two different kinds of listening behavior when they hear the spoken message. As soon as a listener hears the speech which is in the form of acoustic cues, they are translated into the sounds of the language and then into words and phrases and then into an abstract idea. This process is called ‘decoding’ Field (2008). Brown (2011) also defines decoding as ‘a process of breaking up speech into recognizable words’ and these words are tied to background knowledge, systematic knowledge, procedural knowledge and context to form comprehension. In this process, L1 or L2 listener has to make sense of the speech and identify the words. According to Field (2008), only decoding is not enough to create meaning. After creating a literal meaning through decoding, a listener puts contextual or background knowledge on what she has decoded in order to make the speech meaningful. This operation is called ‘meaning building’. In this process, L1 or L2 listener has to apply a different kinds of contextual information, which involves the listener’ knowledge of the word, to the actual words which are produced by the speaker in order to make the speech meaningful.

At a very early level of language learning, gap in the language learner's knowledge of vocabulary or syntax causes decoding problem. The problem may relate to recognition of intonation, sounds, syllables, words or grammatical features. Unlike advanced learners or native speakers, less skilled L2 learners have difficulties in automatic and accurate word matches which creates decoding problem Field (2008). In similar line, Lynch (2009) points out that if language learners have low level of language in listening, it is not an efficient way to encourage them create meaning from context. According to the author, the best thing is to educate them about accurate linguistic decoding and using context. Successful decoding, which relieves listener's attention for meaning building, is a crucial need for L2 learners. Thus, it should be placed emphasis on early in listening programs (Field, 2008).

At a later stages of language learning, weaknesses in learner's listening skill may create meaning building problem. In this stage, a learner may hear a word or grammar part and recognize intonation and sounds but she cannot realize them in natural speech (Ur, 2009). She also cannot handle the information she has elicited from the text efficiently. Furthermore, Field (2008) clarifies the major difference between unskilled and skilled listener is that the novice spends huge amount of effort to the word matching process whereas the expert administers decoding routines automatically, accurately, rapidly and effortlessly. Ur (2009) also states that total familiarity to words and mastering them are a matter of time and practice. Field (2008) also concludes that listeners require to use both input and context but "...using contextual knowledge is more central to successful L2 listening than recognizing words and phrases accurately." (p. 127). As Vandergrift (2000) states that although the cognitive process of listening has been understood more, L2 listening is still the least researched area due to its inherent nature.

2.3.2. Models of Listening Process

We acquire our first language by listening to it even when we are in our mother's womb. After we are born and throughout our lives, listening starts to take up an important part of our communication life and it becomes the most used skill compared to the other ones. (Flowerdew & Miller, 2013). According to Burley-Allen (1995), we spend 9 percent of our daily communication time for writing, 16 percent for reading, 35 percent for speaking, and 40 percent for listening (as cited in Flowerdew & Miller, 2013). Moreover, Mendelsohn (1994) increases the time spent on listening up to 50%. Since listening skill has an important place in our life, many different type of models have been developed to describe how the listening process carries on or how listening comprehension occurs even though there is no commonly accepted one. Besides, several reviews of research (Lynch, 1998; Mendelsohn, 1998; Oxford, 1993; Rubin, 1994; Vandergrift, 2007) have been carried out in order to remark the importance of cognitive models in L2 listening comprehension. The cognitive models of listening such as the 'bottom-up model', the 'top-down model' and the 'interactive model' are the most known of these and they have been used past decades to teach listening. The main aim of the most models is try to explain listening comprehension (Flowerdew & Miller, 2013).

2.3.2.1. The Bottom-up Model

According to the definition for Field (2008), 'bottom-up' explains a process that builds smaller units (sounds) into larger ones (sentences). Flowerdew and Miller (2013) state that listeners build meaning by starting with the individual sounds of the acoustic message and then combine them into words. According to scholars, this process starts with lower level decoding and represented in working memory and then

interpreted in relation to knowledge of context and the world (Flowerdew, 1994; Vandergrift, 2004). Nunan (1999) mentions that this process of decoding is a linear one, in which meaning is created as the last phase of the process. This is a kind of mechanical process in which meaning is built from phonemes to full sentences. In this decoding process, L2 listeners built meaning based on mainly phonological, lexical and syntactic knowledge. (Vandergrift and Goh, 2012). Several studies show that while native or high proficiency level listeners perform bottom-up processing automatically, low proficiency level learners generally do it with much conscious attention to individual units and focus on details of verbal message because of having limited linguistic knowledge in order to understand what they hear (Lynch, 1998; Rubin, 1994; Vandergrift, 2004, 2007).

2.3.2.2. The Top-down Model

According to the definition for Field (2008), ‘top-down’ explains a view of listening as a process that uses larger units (word level) in order to identify smaller ones (phonemes). In this process, listeners construct or sometimes reconstruct the meaning by using incoming sounds as clues (Nunan, 1999). Various researchers (e.g., Flowerdew and Miller, 2013; Nunan, 1999; Vandergrift, 2007) also mention that top-down model underlines the use of listeners’ background knowledge in handling the message. L2 listeners who arrive comprehension in top-down way employ context knowledge of the listening event or the topic of the listening event to interpret the incoming message. Moreover, listeners use different types of knowledge such as prior, pragmatic, cultural and discourse knowledge which are stored in listeners’ long term memory in the form of schemata. In this process, L2 listeners begin to listen with expectation about the information coming from the spoken message and they apply appropriate knowledge to comprehend the message (Vandergrift and Goh, 2012). In

other words, in meaning building process, listeners more rely on contextual knowledge than the sound to decode a verbal message but having prior knowledge or context knowledge about the listening event sometimes are not enough to interpret the message correctly (Vandergrift and Goh, 2012).

Flowerdew and Miller (2013) also add that some people prefer to use bottom-up processing while others rely more on top-down processing. When we look at the language level of the group, beginner language learners are likely to rely on bottom-up skills of decoding. Since these listeners are busy with identifying the words in a speech, they almost do not have enough space for top-down processing (Osada, 2004). However, more advanced language learners, who have knowledge of phonology and syntax, spend more time on improving top-down skills without too much effort. As Conrad (1985) mentions that learners rely on more semantic cues when their level of proficiency increases.

According to Bonk (2000), top-down processing somewhat relies on the knowledge gained through bottom-up processing because if language listeners cannot identify a specific number of words through bottom-up way, they will not be able to draw on top-level cues and understand the meaning of the spoken message. Tsui and Fullilove (1998) also mention that in formal test situation, L2 listeners use more bottom-up processing than they use top-down processing, in addition, bottom-up processing would be more important as a discriminator of listening performance than top-down processing. Compare to reading skill, in listening, language learners rely on more top-down processing (Lynch, 2011). Consequently, the degree to which L2 listeners apply one model more than another relies on their aim for listening because listeners who need to have specific details use more bottom-up processing than

listeners who need to have an overview of a specific event (Vandergrift and Goh, 2012).

2.3.2.3.The Interactive Model

According to the interactive model, which was developed by Rumelhart (1975) (as cited in Flowerdew and Miller, 2013), listening requires both bottom-up and top-down processing. Based on his model, phonological, syntactic, semantic and pragmatic knowledge interact all together but it is not clear how. On the other hand, Field (2008) intentionally avoids using the terminologies ‘bottom-up’ and ‘top-down’. He thinks that using these terms causes problems and they are rather confusing in explaining meaning building, which is from sound to syllables to words and to phrases (bottom-up) or which is using context to enhance decoded message (top-down). He explains that bottom-up and top-down models are not contrasting theories of listening because listening is clearly both as Rumelhart (1975) explained. He also emphasizes that “we could not identify the topic of utterance without some minimal decoding, and we could not appreciate the relevance of the utterance without some minimal use of context and co-text. For this reason, psychological descriptions of listening assume that the processes involved are highly interdependent” (p. 133).

As a justification for his explanation, Field (2008) points out Stanovich’s (1980) ‘Interactive Compensatory Hypothesis’ in his book (as cited in Field, 2008). According to this hypothesis, if a listener is successful at decoding the most part of the input and is confident about the accuracy of decoded input, she will less rely on information presented by context and co-text. On the contrary, if a listener is unclear about what she has decoded, she will be more dependent on contextual information. Furthermore, Field (2008) mentions that both decoding and context play an essential

role in L2 listening regardless of the language level of the listener. In similar line, Lynch (2009) states that context -outside information brought by a listener- supply a link between bottom up (linguistic) and top down (non-linguistic) information. Tsui and Fullilove (1998) differentiate novice listener from experienced one is that novice ones rely more on contextual information to reinforce their decoding because these unskilled listeners have to compensate for gaps in their understanding in which decoding has broken down (as cited in Field, 2008). Moreover, Brown (2011) points out that skilled listeners need both bottom-up and top-down resources but many of the language learners in our classrooms lack ability in bottom-up skill because they are not fluent in listening. According to the author, although these listeners are poor at bottom-up processing, they prefer to use them in listening because they probably think that comprehension is a result of understanding each word. With regard to the familiarity of the content of the listening text, listeners who know something about the content of the text employ their background knowledge to make sense the verbal information, on the other hand, listeners who are unfamiliar to the content of the passage only rely on their linguistic knowledge (Gilakjani and Ahmadi, 2011).

According to Flowerdew (1994), most scholars believe that comprehension include different process, top-down or bottom-up, all of which interact. In similar line, Vandergrift and Goh (2012) define listening as complex cognitive skill and add that both bottom-up and top-down processes are the main part of comprehension. Besides, Voss (1984) states that using of both bottom-up and top-down strategies are better than relying on only bottom-up strategies to interpret sentences for native and second language learners (as cited in Hansen and Jensen, 1994). Similarly, Staehr (2009) states that L2 listeners uses various types of knowledge through bottom-up and top-

down process to interpret the spoken message and expresses that the interaction between top-level and bottom-level cues leads to successful listening comprehension.

To sum up, using either bottom-up or top-down processing depends on the L2 listeners' purpose for listening, listeners' language proficiency level, their working memory capacity, their age, and the context of the listening event. Moreover, in reality, these two process hardly work independently (Vandergrift and Goh, 2012) and interactive model involves the combination of these two processes in which verbal message is interpreted simultaneously at different levels (Suvorov, 2013).

2.4. Cognitive Theory of Multimedia Learning

A theory that explains the interaction between textual and visual information in students' comprehension best is Cognitive Theory of Multimedia Learning (CTML) proposed by Mayer (2005). Therefore, this PhD dissertation was conducted based on theoretical perspective and underlying assumptions of this theory.

According to Mayer (2005a), the term "multimedia" brings different meanings to mind, such as watching a presented graphic accompany with verbal input, or watching a video which includes words and sounds together, or watching a PowerPoint slides presented by a speaker, or writing on a board while presenting a lesson supported by printed text. Hence, he defines the term "multimedia" as "presenting both words (such as spoken text or printed text) and pictures (such as illustrations, photos, animation, or video) (Mayer, 2005a, p.2). In addition, "multimedia learning" refers to learners' construction of mental representations of words and pictures. Therefore, Mayer's (2005a) Cognitive Theory of Multimedia Learning focuses on "the process by which people build mental representations from words and pictures" (p.2). The central idea of this theory is that learners try to form meaningful connection between words and

visuals and they learn better from words and visuals than from words alone. CTML covers essential features of various cognitive theories including Paivio's (1986) dual-coding theory, Baddeley's (1986, 1999) model of working memory, and Sweller's (1999, 2003) cognitive load theory (as cited in Mayer, 2005b).

In order to explain how human mind works, this theory is based on three basic assumptions, namely the dual-channel, limited capacity, and active processing. Dual-channel assumption proposes that human information processing system includes two channels, auditory and visual channels. When learners received both visual and auditory information, they begin to process these incoming information through these two different channels. Secondly, limited capacity assumption is that each channel, auditory and visual, has limited capacity, and they can process limited amount of information at a time. Finally, active processing assumption is that learners engage in cognitive processing actively in order to make sense of the incoming information. The active learning involves three important processes; (a) paying attention to relevant words and images in the presented material, (b) organizing selected incoming information, and (c) integrating incoming information with existing knowledge (Mayer, 2005b).

According to Mayer (2010), learners must engage in five cognitive processes in order to provide meaningful learning from words and visuals. These are; (a) selecting relevant images for processing in visual working memory, (b) selecting relevant words for processing in verbal working memory, (c) organizing selected images into a pictorial model, (d) organizing selected words into a verbal model, (e) integrating the verbal and pictorial representations with each other and with prior knowledge (Mayer, 2005b, p. 38). Figure 1.1 shows a diagram of Mayer's cognitive model of multimedia learning.

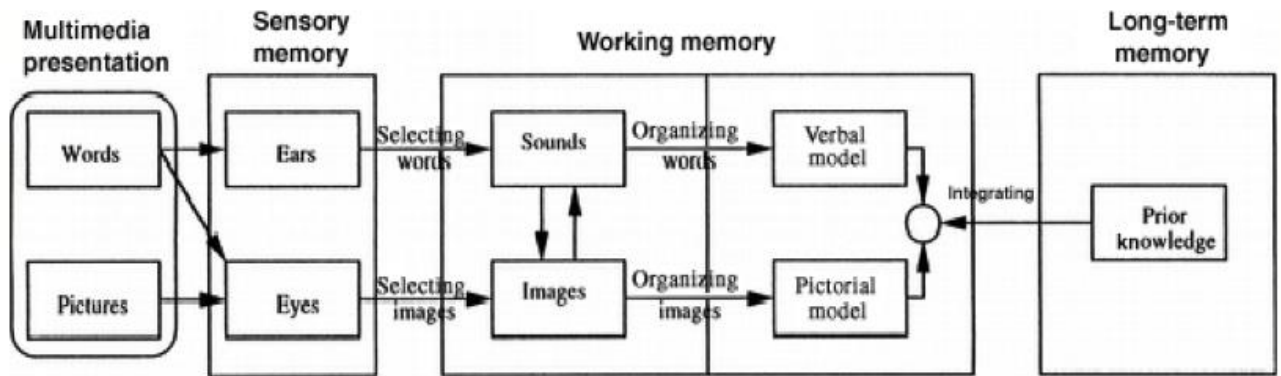


Figure 1.1 Processing of aural and visual information in Cognitive Theory of Multimedia Learning

2.5. Real Life Listening

Teachers in L2 classrooms generally find listening skill complex to teach compared to other skills, reading, writing, and speaking, because listening skill occurs in real time and the spoken message is temporary (Vandergrift and Goh, 2012). In real life, we generally do not listen twice or we do not replay the speech and conversations around us or we do not slow and break the speech down unless we listen to a recorded message or a music or watch a video. Although we ask a speaker to repeat what she said, we never get the same speech but maybe a re-phrasing of it with using of different words. Therefore, in real life, we have only one chance at comprehension (Buck, 2001). Moreover, visual clues around us are of value to help listeners understand what they hear in our real life. In a restaurant, we can see the waiter when he receives our order or in a classroom, we can see our instructor who teaches his subject or in a lecture, we can follow the lecturer's slides about his topic. Ur (2009), in her book, summarizes some features of real life listening as follows: (a) there is an immediate response to what the speaker says, (b) we can see the speaker, (c) the spoken message is supported by visual clues around us. In addition, Lynch (2009) mentions that when listeners begin to listen the first sentences of the speaker,

they have some essential information about the speaker, the place where the speech occurs, the time of the speaking together with visual clues in real life.

Selecting a listening text that reflects real life listening and bringing them into the language classroom is more difficult than selecting a reading text (Buck, 2001). Since most of the audio recordings in language classrooms are not supported by visuals, listening teachers begin to support their lessons by visual materials such as DVDs and videos (Field, 2008).

2.6. Technology in the Language Classrooms

Especially in EFL context, where learners face the foreign language only in classroom environment, it was very difficult for students to practice the target language in outside world. Yet, nowadays, computer technologies and the Internet provide a great variety of possibilities to practice the target language to the language learners. Field (2008) expresses his view on this idea as;

Until recently, it would have seemed hopelessly idealistic to recommend that teachers in many EFL contexts exposed their learners to a wide range of target language speakers. Advances in technology have now provided plenty of material for those who are fortunate enough to have internet connections. But we still have to direct learners to ensure that the material they encounter is relatively accessible both in terms of its language and in terms of its cultural assumption. (p.329)

Parallel to Field's view, Flowerdew and Miller (2013) and Vandergrift and Goh (2012) state that nowadays, most of the language teachers integrate technology such as radios, audio-tapes, videos, DVDs and computers, into their instructions. According to the authors, although radios are low-tech and the least used devices,

they are the most accessible ones in language classrooms. Many benefits of using radios are mentioned in the book of Flowerdew and Miller (2013) as; they provide (a) extensive listening activities to the listeners, (b) native speaker model, (c) news from the world, (d) imagination, (e) accessibility from everywhere, (f) entertainment, and (g) convenience. Today, like radios, with the increasing advancement in technology, audio-tapes become outdated technological tools in language classrooms compared to past ten years. Since the most of the course books came with audio cassettes, audio players were the important devices in our classrooms. The cassettes provided students not only extensive listening activities but also intensive listening practice (Flowerdew and Miller, 2013).

Another important aid which is still actively used in the language teaching is video. Videos provide authentic use of language, paralinguistic elements of the spoken language and cultural background about the language which is used to the language learners (Flowerdew and Miller, 2013). Brown (2011) also mentions that adding visuals to sound, as happens in video recordings, increase learning because of the facilitative effect of visuals on comprehension. In addition, nowadays, increasing the inclusion of video mediated listening activities in listening skill textbooks is the evidence of how videos play a significant role in L2 listening instructions.

Another important and widely used aid in language classrooms today is computers. Warschauer and Healey (1998) express that computers have been used since 1960s in language classrooms and the usage of them are categorized in three areas; behaviorist computer-assisted language learning (CALL), communicative CALL and integrative CALL. In behaviorist approach, language learners perform repetitive drills via computers and all exercises are reading and writing based. In communicative CALL perspective, learners can study a foreign language and do their

language activities on their personal computers in their home with the improvement in technology but activities are still writing based. In integrative CALL today, all language skills, speaking, listening, reading, and writing, are integrated into language instruction via computer technologies. In this trend, huge different kinds of activities and exercises for all skills are available with software and multimedia programs (Warschauer and Healey, 1998). Flowerdew and Miller (2013) and Warschauer and Healey (1998) state the benefits of integrating computer technologies into language instruction as; (1) glossaries are provided, (2) immediate feedback is available for exercises, (3) programs are attractive and fun, (4) learners can customize their learning, (5) all skills are available in one application, (6) programs provide individualization in a large class, (7) pair and small work projects can be organized, (8) real life skill building in computer use (9) programs support exploratory learning. As a justification for the advantages of CALL, Brett (1997) found in his study that multimedia delivered listening comprehension task was more efficient than other delivery modes like audio and video.

Along with the advancement of the Internet, language learners can reach huge numbers of resources for all language skills and fortunately most of them are free and downloadable to personal computers or smart devices via worldwide webs (www) today.

2.7. Academic Listening

Because of large amount of students who want to pursue their studies in English, L2 listening research has been focusing on academic listening (Lynch, 2006). Academic listening means either one way (transactional) listening such as lectures (Flowerdew, 1994) or two-way (interactional) listening such as dialogue between an

instructor and a student in an academic environment (Lynch, 2011). In their book, Flowerdew and Miller (2013) emphasize that listening texts used in the second language classroom are generally similar to a lecture which gives large information to the learners. However, in real life context, the spoken messages are mostly in form of conversations, which are often ignored in our language classrooms. Brown and Yule (1983) describe that lectures are transactional (giving information) which have long turns, whereas conversations are interactional (provide social contact) which have short turns (as cited in Flowerdew and Miller, 2013). Therefore, several researchers (e.g., Flowerdew, 1994; Vandergrift, 2007) separate examining academic listening comprehension from communicative listening in their studies.

Richard (1983) was the first scholar who separated listening skills needed for conversational listening and for academic listening (as cited in Flowerdew, 1994). Richard (1983) mentions the differences in degree and in kind. The differences which are associated with degree as follows: (a) the type of background knowledge: In academic listening, listeners need to know something about the topic of the lecture to understand it easily whereas in conversational listening, there is no necessary background knowledge to get the message, (b) ability to recognize what is relevant to main purpose or what is less relevant, (c) application of the turn-taking conversations, (d) the amount of implied meaning and indirect speech acts. The author also states the differences between academic listening and conversational listening in kind as: (a) necessity of specific skills related to academic listening such as listening to long talks or negotiating meaning, (b) note-taking in academic listening, (c) integrating the message with information coming from other media such as handouts, visuals, or books (as cited in Flowerdew, 1994).

Although presence of some features such as PowerPoint, illustrations, pictures, video texts, or maps in academic listening are important and they facilitate comprehension (see in Gruba, 2004; Sueyoshi and Hardison, 2005), some scholars in research studies identify the difficulties that students encounter in listening academic lectures as speed of delivery, load of new terminology and concepts, cultural differences, note-taking skills, and difficulties in concentrating (Flowerdew and Miller, 1992).

2.8. Testing Listening

Testing language ability is the major element of any language teaching and it is done by asking language learners different question types based on the learnt materials (Flowerdew and Miller, 2013). Whereas it is not so easy to do if we check learners' listening comprehension because as Field (2008) points out that there is no tangible product in listening. Lynch (2011) also mentions that it is difficult and problematic to assess improvement in our listening ability in another language. Flowerdew and Miller (2013, p. 209) conclude their book with these statements that “developing valid and reliable listening language tests is a complex process. This is because the process of listening is hidden from the tester, and so the ways to measure the ability to handle spoken text are more demanding”. Yet a lot of research studies have been done since mid-1990s in order to provide language testers support in designing listening tests (e.g., Buck and Tatsuoka, 1998; Buck, 2001; Freedle and Kostin, 1999).

According to Flowerdew and Miller (2013), teachers generally confuse teaching listening with testing listening because most of them direct their attention to how learners will be successful in listening tests rather than how they improve their

listening skills. The authors also express that in the past language teaching methods, teachers asked learners direct questions in order to reveal how much information the learner could remember, which is called product approach (Figure 1). In this approach, after listening text was introduced to the listeners, teachers check their understanding through questions. If learners could not give the correct answer to the question, the spoken text replayed again. This model is defined more like a ‘test’ by Flowerdew and Miller (2013).

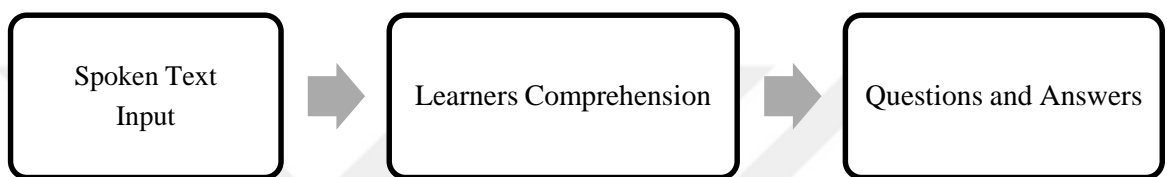


Figure 1. 2 A product approach to using questions in listening

In the book of Flowerdew and Miller (2013), the authors also mention another models for integrating questions while listening, the process approach (Figure 2). In this model, teachers give listeners reasons to listen by providing focus questions or strategies before playing the text and then listeners listen the spoken text and focus on the message which is going to be a part of the post-listening questions.

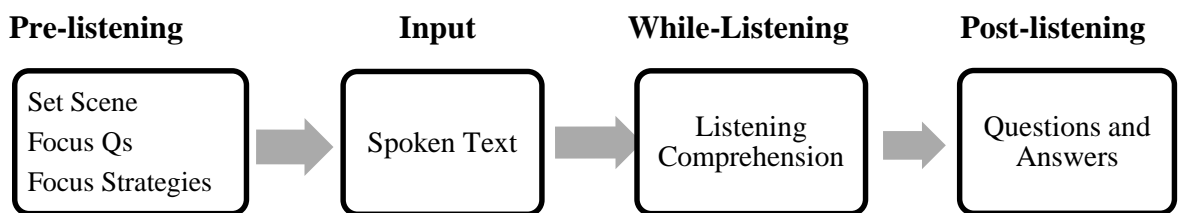


Figure 1.3 A process approach to using questions in listening

Although the previous model reminds us a ‘test’, the second model is like a way which help learners improve their skills and strategies for better comprehension.

Thompson (1995) (as cited in Flowerdew and Miller, 2013) expresses some elements to mark when design a listening test as follows: (a) closer the listening text to oral not written, (b) visual support is important for novice learners, (c) listening texts should be two or three minutes long, (d) a learner can use their prior knowledge to comprehend the spoken message, (e) macro-makers (e.g., Today I am going to talk about...) help listeners understand the text better, (f) very fast speech rate must be eliminated (above 200 wpm).

In his book, Lynch (2011) discusses three main difficulties to assessing second language learners' listening skills. These are: the inaccessibility of mental process, the difficulty of isolating listening skills from the other language skills and other type of knowledge, and test anxiety. In considering the inaccessibility of mental process, Lynch (2011) states that what happens in a listener's mind, her thoughts, and her brain activity when she listens are still impossible to observe. Therefore, researchers and testers have to find another way to assess language learners' listening ability by asking questions, recalling propositions, and identifying pictures (Thompson, 1995, as cited in Lynch, 2011). Yet the designing of a listening comprehension test that reflects the aim of real life listening is still a demanding process.

The second obstacle to assess listening skill, which is mentioned in Lynch's book, is isolating the listening component. The author finds this problem very important because isolating listening skill from other cognitive skills, such as reading or writing, in listening comprehension test is a challenging process. Brown and Yule (1983) suggest the following solutions to cope with this problem in their book: (a) reducing the effect of reading and keeping the questions simple, (b) reducing the effect of writing, (c) reducing the effect of lack of attention, (d) reducing the memory load, using while listening questions, and short listening texts (as cited in Lynch,

2011). The last barrier that testers meet in designing a listening test is ‘test anxiety’. Brindley (1998) mentions the ways that decrease the level of stress in taking a listening test in his article as: making the test instructions clear, allowing enough time for listeners to get familiar with the test format, and involving sample items as a preview.

Researchers have been studying on variables that influence learners’ test performance and try to classify them for many years. Brindley (1998) states some of them in his study as: Nature of the listening input such as speech rate, length of the text, syntax, vocabulary, noise, accent, register, information density, amount of redundancy; Nature of the assessment task such as amount of context provided, clarity of the instructions, availability of question preview, respond required; and individual listener factors such as memory, interest, background knowledge, and motivation. Since these variables effect L2 and EFL students’ listening comprehension performance especially in a listening test, some of these, such as listeners’ background knowledge, rate of the speech, accent of the speaker, the length of the text, text types, listeners’ language proficiency level, presence of the visuals, or test task type, are presented in sub-chapters below.

2.8.1. Background Knowledge

Vandergrift and Goh (2012) define background knowledge as “all conceptual knowledge and life experiences that language learners have acquired and are available for comprehension purposes” (p.65). As stated in the book of Field (2008), the spoken message may be constructed completely different in the listener’s mind. At the end of the listening, the received messages may indicate the listener’s goal rather than those of the speaker. These results happen because literal meaning, which is gained through

decoding, is supported by world knowledge or topic knowledge of the listener. In similar line, Vandergrift and Goh (2012) mention that novice L2 listeners sometimes make incorrect interpretation because of wrong match between their limited linguistic knowledge and world knowledge. Field (2008) defines world knowledge as outside information or background information that makes the speech meaningful for a listener. In similar line, Lynch (2009) mentions that ‘knowledge of the world’, ‘general knowledge’, and ‘non-linguistic knowledge’ are also used to refer the term ‘background knowledge’ in literature. Wu (1998) mentions that listeners use their non-linguistic knowledge when their linguistic knowledge fails them in order to comprehend the speech. On the other hand, Faerch and Kasper (1986) state that “comprehension takes place when input and knowledge are matched against each other” (p: 264). Whatever it is called, the research indicates that activating background knowledge brings about benefit in listening comprehension (e.g., Brown and Yule, 1983; Chiang and Dunkel, 1992; Long, 1990; Markham and Latham, 1987; Park, 2004; Schmidt-Rinehart; 1994; Tyler, 2001).

In terms of research into effects of background knowledge on second and foreign language comprehension, a lot of works have been done on reading skills rather than listening and results of them show that readers find the familiar topic easier than a less familiar one (Lynch, 2009). Moreover, Hansen and Jensen (1994) state that providing context of the listening material by telling what topic listeners will be listening to activates the listener’s schemas. Activating schemas help listeners make predictions about the content and the structure of the recording based on their background knowledge. According to Vandergrift and Goh (2012), if the context of the listening text is not given to L2 listeners at the beginning of the listening for correct interpretation of the spoken text, listeners rely on only bottom-up processing

because background knowledge sources about the text cannot activate suitable schemata. Whereas, when the appropriate context information is provided to the listener, she can apply top-down processing by using her world knowledge sources. The authors also state that activating background knowledge is very important in listening comprehension especially when teaching L2 to adult learners because compared to young L2 learners, adult learners have more life experience (Vandergrift and Goh, 2012). Thus, they can easily make connection between the verbal information and their prior knowledge.

Lynch (2009) concludes that using background knowledge in listening can both support and distort listeners' comprehension. As Vandergrift (2003) states that knowing a lot of about topic of the listening may cause wrong interpretation if listeners cannot check what they understand against the information coming from the input. This result may be attributed to life experience of the listeners because each listener focuses on different points because of the combination of their previous experience and interests (Brown, 1990). Whether listeners' background knowledge facilitate or debilitate listening comprehension, the findings of the studies show that it certainly effects comprehension.

2.8.2. Text Characteristics

Choosing the appropriate text for listening is an important issue both for L2 instruction and L2 testing because it is also a factor that affects listening comprehension (Rubin, 1994). Vandergrift and Goh (2012) point out that many texts are written for reading in language classes. Reading texts include more content words (nouns or adjectives) however, listening texts involve less content words but more function words (prepositions or articles). Since the nature of the spoken text is

different from the written text for reading, there must be criteria based on the selection of the listening texts. Vandergrift and Goh (2012) state seven questions which should be taken into consideration while selecting the texts in their book. These are: (1) What is the original communicative purpose for the material? (2) Who is the intended audience? (3) Who is speaking? (4) What kind of visual support is available? (5) Is the level of the language appropriate? (6) Is the length (duration) of the text appropriate and realistic for the learners? (7) Is the text really meant for listening?

In the book by Field (2008), characteristics of a listening text explains as follows: (1) the text has to be long enough, so at least eight comprehension questions can be created, (2) two comprehension items in the text must not appear closely, (3) medium length, around 3 minutes, recordings are suggested, (4) these recordings must provide full information, (5) single speaker or two speakers from different gender is the best for distinguishing voices (6) as a secret agreement between students and testers, the items in the listening text follow the same order as the points in the text.

In language classroom environment, different kinds of text type should be introduced to L2 learners because the familiarity to the form of the different text types increases students' ability to use suitable skills and strategies while listening (Vandergrift and Goh, 2012). Furthermore, as Shohamy and Inbar (1991) state that the text which is more listenable is found easier by listeners.

2.8.3. Speech Rate and Accent

Lynch (2009) points out that speed in the language materials is the one of the sources of difficulty that language learners face with in both listening and reading classes. Brown (2011), in his book, also mentions that language learners commonly believe that fast spoken message is more difficult to understand than message spoken

slowly because breaking down the speech and realizing the words take time for learners. Therefore, most of the language teachers try to adapt the rate of the speech in recordings as learners increase in language proficiency. For example, in order to slow down the rate, teachers add restatements, synonym, and simple vocabulary or non-verbal signals into the spoken text (Brown, 2011).

As mentioned in Alderson's (1984) article, a reader needs level of 3000 words (Threshold Level) as a necessary requirement for understanding L2 reading. So, a reader can understand a high proportion of the written materials (95% to 97% level). Parallel to Alderson, Nation (2006) mentions that language learners need to know 95 percent of words for reasonable comprehension of a spoken text, (98% is ideal coverage). On the other hand, some researchers (e.g., Kelly, 1991 as cited in Field, 2008) claims that 'Threshold' level for vocabulary is very important in understanding the spoken message but the problems of understanding L2 listening do not occur because of unknown words but because of known words which are not recognized by the listener. This problem may occur in L2 listening, as mentioned in the book of Field (2008), because of the individual characteristics of the speaker such as speaking very fast, slow, loudly, quietly, or pausing frequently. Although different variables such as background knowledge, text type, or language proficiency level of the listeners affect the relationship between speech rate and L2 listening comprehension, various research studies conclude that decreased speech rate increase listening comprehension (e.g., Griffiths, 1992; Jensen and Vinther, 2003; Kelch, 1985; Zhao, 1997).

In order to eliminate the effect of speech rate in test situation, Field (2008) suggests that during the test, comprehension questions should occur after 10-15 seconds of a recording in order to give L2 listeners time to adjust their ears to the

voice of the speaker. Another suggestion made by the author is that playing a recording twice to normalize to the L2 voice.

Another important issue that foreign language learners have to confront with in learning and testing is the accent of speakers. As mentioned in Lynch's (2009) book, there is no evidence that specific accents of English are more difficult than others for language learners. He also expresses that if a language listener is more familiar with a specific accent, she will find that accent less difficult to understand. Furthermore, Hasan (2000) states that if EFL listeners used to listening to specific non-native speaker accent, they may get used to other speakers of different accent but this problem can be overcome through repeated listening and practicing. Lynch (2009) mentions the importance of 'familiarity' to a highly marked accents in order to eliminate the difficulties that foreign language listeners meet. Besides Flowerdew (1994), in his book, points out that accent is generally seen as a cause that creates problem in understanding lectures for non-native listeners. Therefore, the author suggests that the developing of the familiarity with different accent should help learners comprehend the messages easily. Another important point which is mentioned in the book of Ur's (2009) that in today's world, two speakers whose native languages are different may easily communicate each other in English. Similarly, Taylor & Geranpayeh (2011) indicate that providing different accent is desirable especially for students who are pursuing their study in an academic context in which different range of accented information coming from professors or students. Therefore, language teachers may provide learners different accents in listening texts in order to realize different options and to deal with them successfully.

Since speech rate and accent have been mentioned as the significant factors that affecting listening comprehension by second and foreign language learners and

teachers, several studies have been designed to investigate the effect of these two factors (Brown, 2011).

2.8.4. Task Type

Ur (2009) mentions that listening activities will be more useful when they are presented through a task and she points out that “the students are required to do something in response to what they hear that will demonstrate their understanding” (p. 25). This can be any for such as answering questions, naming a picture, or taking notes. The author also states that in a listening activity, tasks may look very charming on paper but they may not work in practice because of some neglected classroom practice such as technical problems, outside noise or students’ low concentration.

According to Vandergrift and Goh (2012), authentic tasks which are designed based on language learners’ age, language proficiency level, and life experience will cause the best listening practice. Moreover, since there is an important relationship between learning and assessment, the tasks which are used in language learning must be the similar ones which are used in the assessment. The authors explain this situation as since language learners learn something that is related to their needs and the learnt information can be used in real life situation, they find the authentic assessment more motivating. Furthermore, this close relationship between learning and assessment task is valid only in achievement tests, which assess the progress in a course. On the other hand, since proficiency and standardized tests do not assess learning in a specific program, content and assessment tasks match global description of listening ability. The authors in their book also state that these tests cannot assess everything in the name of general listening ability. Therefore, testers must put

representative number of language tasks to provide validity (Vandergrift and Goh, 2012).

According to Field (2008), checking listeners' comprehension about listening material through conventional questions is difficult because if students answer the questions incorrectly, it may not be because of failure of listening material, it may be because of the questions which require a lot of reading and writing. If listeners have difficulty in reading (maybe they do not understand what they read) or writing (maybe they cannot formulate a written answer), they may give wrong answers to the comprehension questions. Therefore, listeners' comprehension check should be done through simple tasks which involve grids or filling forms rather than long written questions.

Another important issue that must be taken into consideration is the type of the questions in the tests. Testers or teachers provide either 'display question' - teachers already know the answer of the questions such as yes/no questions or multiple choice questions - or 'referential questions' - teachers do not know definite answers such as summarizing or retelling. Flowerdew and Miller (2013) summarize the reasons of asking display questions as; (1) to ease the listening process, (2) to check basic understanding before facing the more complex issues, (3) to motivate less skilled students to participate the lessons. In addition, Brown (2011) and Hansen and Jensen (1994) in their studies mention another important point that showing questions before listening the audio lead to listen more effectively in the classroom. Similarly, Yanagava and Green (2008) advise that multiple choice questions should be previewed before the listening comprehension test. Thus listeners can make a plan about how to approach the text and which metacognitive strategies can be used.

Hansen and Jensen (1994) explain what listeners must do in answering multiple-choice items as: (a) recognize and store any important information from a mini-lecture as it is read, (b) listen to a question, (c) read four responses, (d) refer back to the information stored in long-term memory to find the propositional information that answers the question, (e) select the response that most closely matches the stored proposition, (f) repeat the whole process for each question. Despite being used commonly- especially in lecture listening- in university content classes, the authors mention the disadvantages of the multiple-choice questions in their study. According to the authors, this type of questions causes cognitive load and if listeners have good memories, read the questions fast or are good at grammar, they will be successful in answering the multiple-choice questions. Hansen and Jensen point out that since multiple choice questions generally require a lot of reading, the listening tests which comprise multiple-choice items do not test listening but test listeners' ability to remember the information. On the other hand, in the study, Yi'an (1998) found that while multiple choice type of questions favor advance listeners, they may be a problem for less able listeners.

2.8.5. Language Proficiency Level

A listening comprehension test involves the interaction of listening stimuli and test questions (Shohamy and Inbar, 1991). Therefore, determining the appropriate listening text, task, and question type, which are appropriate to language proficiency level of the listeners, is one of the most important parts of teaching and testing L2 listening comprehension. Vandergrift and Goh (2012) suggest that the text used in language classrooms should be at a level that learners can comprehend at least at a global level. The authors also add that language learners can control a difficult text if it is listened for global understanding and if the text is provided with a manageable

task. Furthermore, Vandergrift and Goh (2012) remind that testers or teachers should stay away from listening texts which require more background knowledge, involve more unfamiliar words, or produced by unfamiliar accent. Although several studies inform that cognitive processing differs in terms of language knowledge of the learners, which knowledge (i.e., vocabulary, syntactic, background, or discourse) is important at different proficiency level is not clear (Rubin,1994).

As the students' level of vocabulary and syntactic knowledge increases, the students' language proficiency level will increase and their proficiency level will be another important factor in L2 listening comprehension. According to Vandergrift and Goh (2012), vocabulary knowledge and syntactic knowledge are the most crucial elements of successful listening and they play an important role in comprehension for L2 learners. Based on the study of Mecartty (2000), both vocabulary knowledge and syntactic knowledge support reading and listening comprehension. Whereas, the study shows that vocabulary knowledge is more important for reading than it is for listening comprehension. Moreover, the correlation between syntactic knowledge and L2 listening comprehension was not strong enough. In a similar study, Bonk (2000) investigated the interaction between lexical knowledge and listening comprehension in a second language. Fifty- nine Japanese learners of English at different level of L2 proficiency participated to the study. In the study, students listened to four texts of different lexical complexity, wrote recalls and took dictation of the text. Results showed that there was a positive correlation between lexical knowledge and listening comprehension.

Another study which investigates relationship between lexical knowledge and listening comprehension belongs to Staehr (2009). Staehr worked with 115 advanced Danish learners of English as a foreign language in the study. The participants took

three tests in the study. These are a listening comprehension test, a vocabulary size test, and a depth of vocabulary knowledge test. Results showed that there was a strong relationship between students' vocabulary knowledge and the quality of their listening comprehension. Similarly, Vandergrift (2006) mentions that L2 proficiency level, especially vocabulary knowledge, facilitates listening comprehension. To sum up, it is clear that language proficiency is a variable that should be taken into account in every listening comprehension study (Rubin, 1994).

2.8.6. Visuals

Another knowledge type that is used for the listener in order to make the spoken message meaningful is 'knowledge of the setting' as called in the book of Field (2008). This knowledge represents visual cues, which provide contextual framework, in listening and this contextual framework increases comprehension through drawing inferences and monitoring understanding. Vandergrift and Goh (2012) state that the integration of visuals into listening instruction improves the authenticity of listening in the classroom because we already use visuals in real-life listening context and visuals in listening practice is an essential element of listening.

According to Flowerdew and Miller (2013), visual cues are watching the speaker talk or looking at textual information supports and these supports during the listening activity help language learners understand the spoken text better and answer the questions efficiently. Lynch (2011) also states that spoken message which is supported by visuals (face to face or face to screen) directly affect what we comprehend. Moreover, Ur (2009) and Vandergrift and Goh (2012) mention that in the listening activities in language classrooms, visual cues are commonly represented by materials such as diagrams, pictures, sketches, objects, or maps and the existence

of these materials in listening activities provides great benefit in aiding comprehension and bringing to life in classroom. In spite of pointing out the benefits of visuals in language classroom, Ur (2009) also states that although there is the presence of the speaker in real life listening situation, the absence of the speaker in foreign language classroom push the learners concentrated more on the aural message. Vandergrift and Goh (2012) express the drawback of using visuals which occurs especially watching TV news reports in their book. According to authors, what is watching on the TV screen sometimes does not match with what is said on it. Therefore, this incongruity between sound and visual may cause confusion for less-skilled language learners.

In the book of Ur (2009), the author categorizes the exercise types as listening for perception, which main aim is to identify the different sound, stress, and intonation, and listening for comprehension, where listeners make little or no response including other language skills. In the former one, since the focus is on aural perception, teachers or testers do not use visual clues- or keep them minimum- in the exercises.

2.8.7. The Length of Texts

Another important factor that must be taken into consideration while designing a listening test is the length of the listening texts. According to Vandergrift and Goh (2012), the aim of the listening and the intended listening outcomes identify the length of the text. The authors state that if the purpose of the listening is to get detail information, providing a long text can be tiring and break listeners' attention during listening. On the other hand, very short text may cause different problems according to the authors. Vandergrift and Goh (2012) point out that some L2 listeners try to

adjust their ears to sound and the topic of the text when the recordings starts to play. If the listening texts is short, listeners may be lost and the sound will be end until listeners tune in themselves.

2.8.8. Authentic recordings

In many foreign or second language listening classrooms, listeners meet with materials specifically created for language learning. Although these materials help students see various forms and systems of language and how they are used, bringing the authentic materials, which are real life materials, are very essential for language learning process as well (Nunan, 1999). The author also believes that providing authentic recordings to language listeners increases their interest and makes language learning more meaningful. In similar line, Brown (2011) states that authentic materials enhance language learners' motivation due to face validity. The main purpose of the listening instruction is to help learners to understand L2 in everyday situation and this can be done best with authentic materials because they reflect real-life listening in L2 context (Vandergrift, 2007). Although the use of authentic materials in listening instruction is preferred and ideal situation for students in all language proficiency levels, the most important constrain is not to find really authentic texts suitable for low-proficient listeners (Richards, 2006).

According to the book of Field (2008), there is a conservative belief that authentic recordings must be used for more advanced and skilled learners in language classrooms. Moreover, there is another belief that in the beginning stages of the language learning, teachers must use only graded (simplified) materials before listeners are ready for authentic ones. In similar line, Vandergrift and Goh (2012) state that since these materials are created at a normal speech rate, speeches are

clearer and better structured, L2 listeners can easily handle the message by using strategies. However, Field (2008) believes that low-proficient listeners' L2 listening experience should not be limited to scripted materials and we should provide opportunities to hear real L2 aural input. Otherwise, these listeners cannot cope with real L2 listening if they face with it in the outside world. The author also suggests that rather simplifying the authentic text itself, teachers may simplify the tasks which the listeners are likely to achieve, even if the content of the material is above the listeners' knowledge of the language. In similar line, Brown (2011) offers grading the task, not the text. He also points out that language teachers can use any kind of materials, no matter how demanding they are so long as they adjust the tasks according to the proficiency level of the learners.

Vandergrift and Goh (2012), in their book, explain why authentic materials are intrinsically interesting as: (1) the information in authentic materials generally include contemporary topics, (2) these materials involve different types of language use and variety of speakers whom language learners can meet in real life. On the contrary, they also mention that authentic materials may create problems to language learners. Inclusion of hesitation, fast speech rate, accent, and pauses in the natural speech can be challenging situation for L2 learners. Therefore, the authors suggest to use scripted or semi-authentic materials in the language classrooms.

Ur (2009) mentions two disadvantages of using authentic recordings in foreign language classrooms. (1) authentic recordings are generally suitable for advance level of learners because the language of them are not graded, so it is difficult, (2) natural conversations are hard to understand even for native speakers without seeing the speaker, thus authentic sound recordings have minimal value in foreign language classrooms. Furthermore, Ur (2009) does not believe that authentic materials provide

real life listening. As a justification for this claim, Lynch (2009) mentions that many programs and speeches from BBC World Service, CNN News or United Nations are not authentic because their target audiences are generally non-native listeners. In addition to these views, Nunan (1999) expresses that since it is difficult to understand authentic recordings, they provide threats to language listeners.

When opinions of allies and opponents on using authentic materials are considered, Brown (2011, p. 135) summarizes their ideas as: (1) The strong authenticity position: language is best learned if all input is authentic, (2) the non-authenticity position: language is best learned if all input is specially written for the learners, (3) the intermediate authenticity position: language is best learned if input is varied in degree of authenticity according to the learners' proficiency and the purpose of the lesson at that point in the curriculum. The author also concludes that language learners should see all types of input whether it is authentic or not. Therefore, providing authentic materials sometimes is useful and sometimes it is not. McGrath (2002) provides important principles for choosing authentic recordings for language learners. These are (a) relevance (to syllabus, to learners' need), (b) intrinsic interest of topic/ theme, (c) cultural appropriateness, (d) linguistic demands, (e) cognitive demands, (f) logistical considerations (length, audibility), (g) quality, (h) exploitability (as cited in Field, 2008).

2.9. Role of Visuals in Second and Foreign Language Listening Comprehension

Field (2008) mentions that L2 listeners, especially novice ones, trust the information coming from external world rather than relying on the evidence of their ears when they listen. The author explains this situation in two ways: (1) listeners feel themselves so unskilled that they do not expect to discriminate sounds in the second

language. They do not trust their own abilities. (2) Since novice listeners cannot decode automatically, decoding process makes huge demands upon attention. Thus, less experienced listeners choose to rely upon more external messages which are less challenging for the listeners. Furthermore, Ur (2009) expresses that since visual materials in language listening activities attract listeners' attention, these aids encourage them to focus on the listening texts more.

As it is explained in the book of Flowerdew and Miller (2013), listeners need four types of knowledge in order to comprehend spoken message. These are: 'phonological'-the sound system; 'syntactic'- how words and sentences are constructed; 'semantic'- meanings of the words and sentences; and 'pragmatics'- meaning of language in particular situations. Moreover, the authors suggest another type of knowledge which gives listener contextual clues. It is 'kinesic' knowledge. Having a knowledge about the meaning of non-verbal communication - gestures, mimics, eye contact, posture and body movement- help listener convey the spoken message. According to Flowerdew and Miller (2013), although relying on only audio recordings reduces the exposure of listeners' kinesic knowledge, a lot of reliance on videos and real life listening, where the speaker is visible, increases this essential dimension of meaning. Likewise, Ur (2009) mentions that listening without seeing (heard blind) is very demanding process compare to listen with something, which is colorful or in motion, to look at.

In a similar line, Ur (2009) states that in real life circumstances, the visibility of the speaker, her facial expressions, and movements help listeners understand the spoken message easily. Therefore, the author does not find right not to use visual clues in classroom practice. Addition to Ur's view, Lynch (2009) points out that visuals and non-verbal clues play an important role in comprehending what speakers

say and he suggests that foreign language teachers should integrate them into listening instruction whenever they are possible.

On the contrary to the views of Field (2008) and Flowerdew and Miller (2013), Ur (2009) finds out that several foreign language learners may not use visual clues when they listen a text or learners use the clues but they misunderstand what they hear. She clarifies that language learners confront this difficulty because they either analyze the speech in isolation or they do not link the speech to the visuals (e.g., learners may understand horse for house in spite of having an illustration).

Contrary to the view of Field (2008) who claims that since unskilled foreign language listeners cannot decode automatically as they do in their native language, they depend on more environmental clues to comprehend the spoken message, Ur (2009) declares that language listeners are not able to perceive and interpret visual clues in listening a foreign language in spite of doing so well in their native language. According to the author, the reason why listeners cannot apply environmental clues in their interpretation is that listeners' memory system is so overloaded with decoding every single item coming from the foreign spoken message that listeners cannot relax and find a time to take a broader view.

Vandergrift and Goh (2012) suggest some key points to teachers and students who want to employ visual media for L2 listening based on their research literature in their book as; (1) using visuals in language learning causes positive effect on students but visual and aural input together sometimes put a load on working memory and they may be distracting for listeners, (2) relevant visuals help listeners stimulate their metacognitive knowledge to guess about the context of the text and listeners can use strategies to improve their weak linguistic knowledge, (3) there must be match between the content of the visuals and aural message, especially for less-skilled

language learners, (4) in listening tests, testers must be very cautious about involving the visuals into the test because eye contact between screen and test material may cause negative effect on listeners, (5) media literacy courses should be included in learners' curriculum.

Along with different types of visuals, several researchers have been also investigating the effect of different types of input modes such as audio with visuals and video on listeners' L2 or FL listening comprehension in listening tests (Batty, 2015; Coniam, 2001; Gruba, 1993; Ockey, 2007; Sueyoshi & Hardison, 2005; Suvorov, 2009; Wagner, 2007; Wilberschied & Berman, 2004).

2.9.1. Types of Visuals

In second and foreign language listening research, visuals are distinguished as context and content (Bejar et. al., 2000; Ginther, 2002; Ockey, 2007). Context visuals involve the visuals about the context in which verbal information occurs (Ginther, 2002). This type of visuals is used to present either the image of participants or settings in a listening text. For example, a photo of a professor giving a lecture in a conference hall depicts context visual. In a similar line, Bejar et al. (2000) classify context- or "situational" according to their classification- visuals in terms of information that they present. These are visuals of (a) setting, (b) participants, and (c) text type. Furthermore, in literature, different scholars use different terminologies for visuals that provide information about setting and participants. For example, Ur (2009) defines environmental clues as visuals and adds that "environmental clues are often more likely to provide information about the situation, speakers and general atmosphere than about the actual topic of discourse" (p. 5). Similarly, Vandergrift and Goh (2012) point out that watching a speaker while she is speaking facilitates

listening comprehension as it happens in real-life situation because seeing the speaker provides kinesics cues (body language, facial expressions, and gestures), which facilitate comprehension of the speech.

The second type of visuals that is defined by Ginther (2002) is content visuals. This type of visuals involves a photo, a map, or a graph that illustrate the content of the verbal information. For instance, a photo of junk food accompanying a dietitian who is talking about healthy diet. As stated in the study of Bejar et al. (2000), content visuals are classified in four ways according to their relationship to the verbal information. These are: (a) visuals that replicate oral stimulus; (b) visuals that illustrate the oral stimulus; (c) visuals that organize information in the stimulus; (d) visuals that supplement the oral stimulus. The authors also mention that while the first three types of content visuals facilitate listeners' listening comprehension, the last type make the comprehension more difficult (Bejar et. al., 2000).

Research studies which examine the effects of visuals on listening comprehension generally pertain to visual-non visual distinction (Suvorov, 2013). However, when the applicability of context and content visuals in research designs in the literature is examined, most of the visual types in which the effect of them was investigated was context visuals (e.g., Coniam, 2001; Ockey, 2007; Suvorov, 2009; Wagner, 2007, 2008, 2010a, 2010b). There are only few studies which included the content visuals in their design in the literature. (Ginther, 2002; Suvorov, 2013).

2.9.2. Types of Input Modes

As much as the differences in the types of the visuals (context or content), the input modes to which these visuals are presented are also important in a second and foreign language listening comprehension test. Several types of input modes are

possible that can be applied into listening comprehension test, such as audio-only, series of context or content still images accompany by audio, or context or content visuals accompany by video (Ockey, 2007). Although there are some studies which investigate the effects of visuals through audio with still image input mode (Ginther, 2002; Ockey, 2007; Suvorov, 2009) in the literature, majority of them focus on comparing the effect of visuals delivered via audio-only and video mode (Coniam, 2000; Gruba, 1993, 1997; Londe, 2009; Progosh, 1996; Suvorov, 2009, 2013; Wagner, 2010b).

Research on the use of visuals in listening comprehension suggests that the use of video can promote students' listening skills and is more helpful for less proficient language learners, especially when they encounter difficult text. If teachers consider to use authentic materials in the language classroom, variety of choices are available, especially video recordings (Field, 2008). Ur (2009) also states that authentic sound recordings are not often used in language classroom because the language used in the texts is difficult to understand for the language learners, even for the natives.

Therefore, rather than using authentic sound recordings, authentic video recordings might be better materials for high level language learners. Lynch (2011) mentions that the use of video recordings in a language classroom improves the students' likely meet with the second language. Moreover, Brown (2011) expresses that integrating videos in language teaching increases students' motivation because learners like activities that resemble real life. Therefore, especially in today's world, videos are used as a language instruction tool in almost every language programs.

According to Flowerdew and Miller (2013), the use of video as an instructional tool in language classrooms began in the mid-1970s and from now on it has been used especially in the countries in which English is learnt as a foreign language. As

mentioned in the book of Flowerdew and Miller (2013), watching videos help language learners learn not only authentic use of language but also cultural context of it. Moreover, since videos involve viewing process, they help learners comprehend the spoken message through paralinguistic features. According to these authors, videos can be used as both intensive listening activities in the classroom and extensive listening practice at learners' home. Although watching videos generally can be seen as an entertainment activity in the classroom, teachers can turn them to an instructional practice through tasks, worksheets and follow-up activities. Furthermore, Garis (1997) (as cited in Flowerdew and Miller, 2013) mentions some possible problems of choosing appropriate videos as an instructional tool in the classroom. For the author, teachers must be careful in terms of obscenity, brutality, profanity and other concerns which might embarrass some learners.

Another important issue that must be taken into consideration according to Vandergrift and Goh (2012) is that concurrency of the visual and aural input in video texts. When there is no match between these two elements of the video, language learners, especially novice ones, may lose their motivation, find the assessment task demanding, and finally comprehension process may fail. Thus, main attention in preparing video for listening comprehension is to provide precise combination between the content of the images and the aural message.

2.10. Studies on the Impact of Visuals in Listening Comprehension

According to Vandergrift and Goh (2012), the result of research studies that investigate the facilitative effects of visuals on listening comprehension is contested. The authors claim that although language learners are positively affected by the appearance of visual materials in listening texts, they rarely benefit from them for

comprehending the spoken message. They also mention that, based on the findings of qualitative results, the using of visual materials in listening have positive impact on language learners but how much the visuals facilitate comprehension is not conclusive. Because of the mixed results of the research studies in the literature, the effects of context and content visuals, which are delivered via different types of input modes, on second and foreign language listening comprehension have been investigated for several years (Brett, 1997; Coniam, 2001; Ginther, 2002; Gruba, 1997, 1999; Homayoun, 2013; Ahanghari, Mozaheb and Mirzapour, 2013; Londe, 2009; Maleki and Rad, 2011; Ockey, 2007; Progoosh, 1996; Shin, 1998; Suvorov, 2009, 2013; Wagner, 2010). Among these studies some of them indicating that visuals help listeners perform better on tests (Ahanghari et. al., 2013; Ginther, 2002; Latifi and Mirzaee, 2014; Maleki and Rad, 2011; Mueller, 1980; Wagner, 2010b) while others have found detrimental effect (Suvorov, 2009) or no effect (Coniam, 2001; Londe, 2009; Gruba, 1993; Ockey, 2007; Shin, 1998; Suvorov, 2013).

Suvorov (2009), for instance, investigated three areas in his study. First of all, he examined the effects of context visuals among three different types of modes, which were audio-only, videos, and audio with single photograph. He also investigated the effects of text types- a dialogue and a lecture- on students' L2 listening test performance. Finally, he investigated the participants' preferred type of visuals and their effects on the parts of the listening test. Thus, he designed a computer-based listening test, which involved six listening passage and 30 multiple choice questions with five options. 34 ESL students from high to low English proficiency level at a public university participated in the study. Suvorov used three types of input modes in his test, a single photograph format, video format, and audio-only (no visual) format. In these formats only context visuals are used as a type of visuals. Each format

involved two text types, a dialogue and a lecture. The result of the ANOVA showed that there was no statistically significant difference between ESL students' audio-only scores and audio with single photograph scores. On the other hand, statistically significant difference was found between participants' audio-only and video scores, as well as audio with single photograph and video scores. According to descriptive statistics, while participants received higher scores in audio-only mode, they got lowest scores from video mode. Regarding the investigation of the text types, ANOVA analysis revealed that only the mean score of the video lecture was significantly lower than the mean scores of other five listening passages. Therefore, the researcher concluded that the use of video in lectures had a detrimental effect on test takers' performance. Overall, Suvorov concluded that although the use of audio-only and audio with single photo format did not make any significant difference on students' test scores, video mode made a negative impact on the scores, especially videos in lectures had a detrimental effect.

Another important comparative study of audio supported by a series of context still images and video included context images in academic computer-based listening test was conducted by Ockey (2007). In the study, the researcher investigated how participants engaged with these two input modes. Six university level nonnative speakers of English with three ability levels- advanced, upper-intermediated, and lower-intermediated- took two listening tests: videotaped version and audiotaped with five still context images. Furthermore, interviews, retrospective verbal reports, and videotaping of the test takers were used as data collection instruments. One of the results of Ockey's study showed that although all participants engaged with still images minimally during the test, their engagement with visuals in video was more extensive. The result of the study also showed that still images found to be beneficial

at the beginning of the listening test, because they provided context, but were not helpful later during the test. Although all of the test takers generally did not find the still images distracting, video stimulus was found to be no help and distracting for some of them.

Another important study that investigated the difference between audio and video mode belongs to Gruba (1993). 91 advanced ESL students participated to this study and they took 14-item (multiple choice and true/false) academic lecture listening test. Listeners' performance on L2 listening test was compared on two modes, audio and video. The result of the study indicated that there was no statistically significant difference between audio and video scores of the students. According to the researcher, one of the reasons that created this result was low reliability (.45) of the test. Furthermore, Gruba stated that no difference between two modes was occurred due to the possibility that advanced language learners are not "medium-dependent" (p:87).

In a similar line, Coniam (2001) investigated the effects of audio and video mode in an EFL listening test. Groups of pre-service and in-service English language teachers who enrolled in a post-graduate diploma in Education program in Hong Kong took audio (N=57) and video (N=47) version of the same test. After completing the listening test, participant completed a questionnaire in order to assign their preference for audio or video mode. The result of the t-tests indicated that although audio groups performed slightly better than video groups in the listening test, there was no statistically significant difference between two groups. In addition, the result of the qualitative data showed that although some of the participants in audio group preferred to take the test via video mode, the participants in the video group stated

that they did not gain any advantage from the video and they might have done better in audio mode.

Another important comparative study of audio-only mode and video mode was conducted by Londe (2009). The aim of her research study was to investigate whether there was any significant difference in students' L2 listening comprehension test performance in different formats. 101 ESL students from mid-high to high level English proficiency participated in the study. One-way ANOVA was used to compare two video formats (i.e., close-up view of the head of the lecturer, and full body view of the lecturer and one audio-only format). The result of the analysis showed that regardless of the input format the students in three groups performed similarly. Thus, the researcher concluded that input format did not affect students' listening comprehension.

Contrary to the studies which used only context visuals either in audio with visuals mode or in video mode, Suvorov (2013) investigated the effects of both context and content visuals on ESL students' listening comprehension. The researcher designed a video-based academic listening test (VALT), which comprised six video clips (three of them with context visuals and three of them with content visuals) with 30 multiple-choice items for the study. Each video clip included real-life academic lectures. Students' also took the audio-only version of VALT, which was called AALT. Both VALT and AALT were delivered via Moodle. After administering the VALT (N=75) and the AALT (N=46), students' listening test performance scores were analyzed using pair-sample t-test and independent sample t-test. The analyses of the test performance data showed that visuals did not have any effect on students' L2 listening comprehension.

Unlike the findings of the previous studies mentioned above, several researchers reported facilitative effects of visuals on students' performance on second language listening comprehension tests. For example, Ginther (2002) is another researcher who investigated the effects of both context and content visuals on L2 students' listening comprehension. Ginther, in her study, also examined the effect of type of stimuli (dialogue, short conversations, academic discussions, and mini-talks) and language proficiency level (low or high) on TOEFL CBT listening comprehension test. 160 participants (80 high proficient and 80 low proficient) took 40 item (20 of which were accompanied by visuals and 20 were not) listening comprehension test. In the test, while context visuals were employed in all type of stimuli, content visuals were involved only in mini-talk. The result of the study showed that although content visuals in mini-talks increased participants' comprehension, context visuals accompanied with mini-talks did not boost test takers' listening comprehension. Context visuals had slightly debilitating effect on test-takers' performance. Ginther (2002) also concluded that if the visuals in the audio parts are contiguous and if they are related to the content of audio message, facilitative effect will occur. Furthermore, she suggested that test takers may be distracted by context visuals when the content of the verbal channel is difficult.

Another important study, which compared audio-only mode and video with context visual mode, was conducted by Mueller (1980). 123 low proficient and 76 high proficient level English speaking students of German participated in the study. Context visuals in the video mode involved a simple line drawing of the participants in the interview and their relationship to each other. After watching the videos and listening the audio text, participants asked to write a seven-minute summary of the passage. The result of the study indicated that contextual visuals facilitated beginner

level German students' listening comprehension. The researcher also concluded that the effect of context visuals inversely related to students' language proficiency level. That is, the more proficient learners gain less help from visuals because of their high language skills.

In another research study, Maleki and Rad (2011) investigated the effect of static images to verbal stimuli on an IELTS listening comprehension test. 58 EFL students who had higher and lower level of proficiency participated in the study. Test takers took two different types of modified computer-based listening input. One with involved a series of still images related to the verbal stimuli and other one included textual information consisting of a series of paraphrased script presentation of the same verbal stimuli. The result of the study indicated that both higher and lower proficiency level listeners performed much better in listening tests which involved more static images than students in listening tests which included fewer static images. The researchers concluded that higher frequency of static images accompanied with verbal stimuli enhanced test takers' performance on listening comprehension test. They also stated that visual aids helped lower proficient students more than higher proficient students.

In addition, Wagner (2010b) conducted a study to investigate the effect of videotexts on ESL students L2 listening performance. 202 students with six different language proficiency levels participated in the study. While 103 of them were assigned to experimental group and took the input through video text, 99 of them were assigned to control group and took the same version of the listening test with audio-only text. Total eight video text, which included four dialogues and four lectures, with multiple choice items and short answers items were used in the test. The researcher used Multi-Variate Analysis of Covariance (MANCOVA) to compare the

test performance of the two groups. The result of the analysis showed that video group scored 6.5% higher than the audio-only group and the researcher concluded that visual information in the video texts contributed to the video group's superior performance.

To sum up, in most of the existing research studies, context visuals are the most commonly used visual type and videos are the most commonly studied input mode. As can be seen in above mentioned studies, the researchers generally compared students' L2 or FL listening comprehension scores in audio-only mode test with their scores in video with context visual mode test. However, there is no agreement among researchers on whether visuals (context or content) facilitate students' second or foreign language listening comprehension, or on whether which input modes (audio with visuals or video) make a difference for listeners. Thus, this inconsistent findings of the existing research studies must prompt researchers to conduct more studies in the field of second and foreign language listening comprehension.

CHAPTER 3

METHODOLOGY

The purpose of this chapter is to explain the methodology of the study which includes research questions, the description of research design, setting, participants, and data collection instruments and procedures, and data analysis.

3.1. Research Questions

The aim of the present study is to investigate the impact of context and content visuals in multi-modal inputs (video, audio with visuals, and audio-only) on English as a foreign language students' academic listening comprehension. In particular, the study examines; a) whether there exists a difference in students' performance on five listening subtests, namely, audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals, and b) whether individual differences, that is English language proficiency level, gender, and students' listening style, affect EFL students' academic listening comprehension. Therefore, the study addresses the following research questions to be answered.

1) Is there any significant difference among different types of input modes, namely audio-only, audio with visuals, and video, in the listening tests in terms of their impact on EFL students' academic listening performance?

2) Is there any significant difference between different types of visuals, context and content, in the listening tests in terms of their impact on EFL students' academic listening performance?

2.1) Is there any significant difference between audio with context visuals and audio with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

2.2) Is there any significant difference between video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

2.3) Is there any significant difference among audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

3) Is there any significant effect of individual differences (proficiency level, gender and listening style) on EFL students' performance scores in academic listening subtests, namely audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals?

3.1) Is there any significant effect of proficiency level on EFL students' performance scores in academic listening subtests?

3.2) Is there any significant effect of gender on EFL students' performance scores in academic listening subtests?

3.3) Is there any significant effect of listening style on EFL students' performance scores in academic listening subtests?

4) What are EFL students' experiences about the presence and absence of context and content visuals provided by multi-modal input in the listening tests?

3.2. Research Design

In the present study, causal-comparative research design, which is also known as ex post facto design, was adopted. In this design, researchers "attempt to determine

the cause or consequences of differences that already exist between or among groups of individuals” (Fraenkel & Wallen, 2006, p.370). Although causal-comparative research is similar to an experiment, the independent variables are not manipulated in this design. Researchers compare groups in order to find out the effects of the already occurred independent variables on the dependent variable. In the framework of the design, two listening tests, listening test-1 and listening test-2, involving ten similar listening subtests given to the study group in order to collect data and analyze the differences among them. As can be seen in table 3. 1, table 3. 2, and table 3.3, four sets of data were collected for this study; (a) demographic data, consisting of personal information about the participants of the study, (b) EFL Listening Style data, showing the participants’ listening styles -visual, auditory, spatial, or bottom-up, (c) test performance data from Listening Test-1 and Listening Test-2 (include ten audio and video texts with and without visuals), (d) interview data, consisting of participants’ verbalizations of their opinions about the listening tests.

Table 3.1 Data Sets Collected for the Study

Data Set	Data Type	Data Collection Method	Data Sources
Demographic Data	Quantitative	Demographic Information Questionnaire	n=127
EFL Listening Style Scale Data	Quantitative	EFL Listening Style Scale	n=127
Listening Tests Performance Data	Quantitative	Listening Test-1 Listening Test-2	n=127
Semi-structured Group Interview Data	Qualitative	Interview Questions	n=24

Table. 3.2 The Design of the Study

Listening Test-1							
Study	Demographic	EFL Listening	Audio with	Audio with	Video with	Video with	Audio Only
Group	Questionnaire	Style Scale	Context Visual Part-1	Content Visual Part-2	Context Visual Part-3	Content Visual Part-4	Part-5
N: 127							

Table. 3.3 The Design of the Study

Listening Test-2						
Study	Audio Only	Audio with	Video with	Video with	Audio with	Semi-structured
Group	Part-1	Content Visual Part-2	Context Visual Part-3	Content Visual Part-4	Context Visual Part-5	Group Interview
N: 127						

3.3. Setting

The present study was conducted at the English Language Institute of a private university in Istanbul, Turkey in 2014-2015 academic years. This university established as a thematic university in order to raise social scientists and acts within very limited number of departments. This university is an English medium university which involves two faculties (faculty of management and administrative science and faculty of humanities and social science) and two institutes (institute of social science and English language institute) with approximately 1,600 students.

Since English is the language of instruction in all departments, except in the departments of Turkish Language and Literature, all incoming students whose native language is not English have to take English Language Proficiency Exam, which is designed by English language institute, or present TOEFL or IELTS test scores (score of 550 or above for paper-based TOEFL, 80 or above for IBT TOEFL or 5.5 or above for IELTS) at the beginning of the fall semester. The students who have a score of 70 or above from the proficiency exam are allowed to register to courses from their departments. Yet, the students who take the score of less than 70 have to take the Oxford Online Placement Test, which is designed by Oxford University Press, in order to be placed into their classes according to their level of English at English language institute. After assigning the students' English language levels based on the results of the placement exam, English language teaching starts according to their levels. The aim of the institute is to teach English for the general academic purposes.

In the English language institute, English language is taught as a foreign language in four different language levels, A1(breakthrough or beginner), A2(waystage or elementary), B1(threshold or intermediate) and B2 (vantage or upper-intermediate). These language levels are formed based on the Council of Europe's Common

European Framework of Reference (CEFR). The CEFR involves three main language levels- A, B, and C. These main levels are subdivided into two levels (A1, A2, B1, B2, C1, and C2). A1 is the lowest level and C2 is the highest level of proficiency described in the CEFR (Council of Europe, 2001). In the institute where the data were collected for this study, B2 is the highest level. Students who completed B2 level can take courses from their departments.

At the beginning of the 2014-2015 academic years, 235 students were registered to English language institute in order to learn and improve English. After the results of the placement exam, 180 students were placed into A1 level, 35 students were in A2 level, and 20 students were in B1 level. The number of the students in each class was around 18 and there were 30 English language teachers (5 native speakers of English and 25 non-native speakers) at the institute. Each class had one computer, one smartboard system, and a surround sound system for listening and video watching activities. Students in the institute took 28 hours of English language instruction, which were generally taught by four different native or non-native instructors, in a week at the time of the study.

In this institute, each academic year consists of four modules, one of which takes 7 to 9 weeks long. At the end of each module, students who complete the requirements of each language level (e.g., two achievement and one final tests) successfully can pass to an upper level (e.g., students who are in A1 level will be in A2 level after completing the requirements of A1 level successfully). However, if learners cannot pass to an upper level at the end of each module, they have to repeat the same level that they have already studied. In the present study, data were collected in the second module (between January 5th, 2015 and January 21st, 2015) and in the third module (between March 18th, 2015 and March 27th, 2015) as can be seen in

figure 3.1, the reason why the researcher collected the data in two different modules is to provide maximum participation of the students from three different language levels to the study.

3.4. Participants

The participants of the present study were, one hundred-twenty-seven EFL students- sixty-nine female and fifty-eight male- in English language institute of the private university. At the time of the data collection, majority of the participants were aged between 18 and 20 (85%). Since the culture of which we are members may affect the perception of the listening texts, only Turkish students were included in this study. As can be seen in figure 3.1, the number of the participants who were in A2 level (elementary language level of English) was 42, B1 (intermediate language level of English) was 42, and B2 (upper-intermediate language level of English) was 43 at the time of the data collection process. The EFL students were placed in their English language proficiency levels based on their scores of the computer-adaptive placement test, which was administered online in the computer labs at the beginning of the academic year. Later on, all students start to learn English language in their levels in the first module. Each student' English language proficiency level changes based on his/her average scores of achievement and final tests, which are administered in every 2-month modules. If students' average scores of these achievement and final tests are 70 or above, they will pass to an upper proficiency level. If their average scores of these achievement and final tests are below 70, they will repeat the same level for another 2 months.

In the setting where the researcher collected data, teaching academic level of English starts in A2 level. Since beginner level (A1) students did not have enough

English language background and could not handle academic listening tests because of their insufficient academic English level, the target population of these tests involved students with only A2, B1, and B2 levels of English language proficiency. Therefore, the rationale for including these three English language proficiency levels (i.e., A2, B1, and B2) but not including A1 (beginner) language level of English purposefully in the study was related to the researcher's intention to provide representative sample of the target population in the listening tests used as a data collection instrument in the study.

	A2 LEVEL	B1 LEVEL	B2 LEVEL
2nd MODULE (5th Jan. - 21st Jan.2015)	42 (3 Classes)		30 (2 Classes)
3rd MODULE (18th March- 27th March 2015)		42 (3 Classes)	13 (1 Class)
TOTAL NUMBER	42 A2	42 B1	43 B2

Figure 3. 1. Number of the participants in the study in 2nd and 3rd module.

Table 3.4 presents demographic information about the study participants that was obtained via a demographic information questionnaire. The collected background information involves the participants' age, department, level of English language proficiency, number of years they had been learning English, personal information about watching TV or listening to radio, preferences to improve their listening skill, and amount of understanding in listening English.

Table 3.4 Background Information about the Participants (n=127)

Gender	Age	Faculty	Proficiency Level	Length of studying English	Watching English TV Program	Time spending on watching these program	Amount of Understanding
female (n=68)	under 18 (n= 3)	Humanities and Social Sciences (n=60)	A2 (n=42)	less than 1 year (n=15)	Yes (n=108)	less than 30 min. (n=15)	less than 25% (n=10)
male (n=59)	18-20 (n=108)	Management and Administrative Science (n=67)	B1 (n=42)	1-4 years (n=18)	No (n=19)	30min-1 hour (n=34)	25%-50% (n=56)
	21-23 (n=13)		B2 (n=43)	5-10 years (n=61)		more than 1 hour (n=33)	51%-75% (n=30)
	23+ (n= 3)			10+ years (n=33)		other (n=29)	More than 75% (n=14)

Table 3.4 Background Information about the Participants (n=127)

Listening to Radio	Time spending on listening to radio	Amount of Understanding	The way improving English	Amount of Understanding English Lessons
Yes (n=14)	less than 30 min. (n=8)	less than 25% (n=5)	watching TV (n=55)	less than 25% (n=1)
No (n=111)	30min-1 hour (n=4)	25%-50% (n=2)	listening to radio (n=3)	25%-50% (n=17)
	Other (n=2)	51%-75% (n=7)	listening podcasts (n=13)	51%-75% (n=39)
		More than 75% (n=0)	listening to English lessons (n=22)	more than 75% (n=70)
			listening to native friends (n=28)	
			Others (n=6)	

3.5. Data Collection Instruments

The data of the present study were collected through four different instruments which were developed by the researcher herself. These instruments involve Demographic Information and Information of English Listening Questionnaire (see Appendix A and B), English as Foreign Language Listening Style Scale (see Appendix E and F), Listening Test-1 and Listening Test-2 (see Appendix S and T), semi-structured group interview (see Appendix U). Detailed descriptions of the instruments are given in this chapter.

3.5.1. Demographic Information and Information of English Listening Questionnaire

Demographic Information and Information of English Listening Questionnaire was developed to collect personal information about the participants of the study. It includes 13 questions that asked participants to provide information about their age, gender, English language level, major at the university, the number of years that they had spent studying English, preference for watching or listening in English, ways to improve their English, and amount of understanding English. The original version of the questionnaire was Turkish but it was translated into English for further use. (see Appendix A and B). The collection of the personal information through this questionnaire was accomplished before the listening tests were administrated.

3.5.2. EFL Listening Style Scale

In order to investigate the effect of EFL students' listening style on their performance scores in academic listening subtests (it is the research question 3.3 in the study), the researcher needed to have a scale that reveals EFL learners' listening style when they listen to a text in English language. After investigating various

learning style scales specifically focusing on English language learning (e.g., Kinsella, 1995; O' Brien, 1990; Oxford, 1993; Reid, 1984) and very few listening style inventories (e.g., Bodie & Worthington, 2010; Pearce, Johnson, & Barker, 2003) aiming to determine participants' perceived listening effectiveness when they are communicating in their native language in the literature, it was surprising that no questionnaire aiming to determine foreign or second language learners' listening style when they listen to a material in English language was available in the field. Therefore, English as Foreign Language Listening Style Scale (EFL-LSS) was decided to be developed in order to fill the gap in the field.

3.5.2.1. Developing the EFL Listening Style Scale

EFL Listening Style Scale, which has 17- items, was developed to reveal EFL language learners' listening style when they listen a material which was produced in English. Particularly, it helps EFL language learners better understand their internally based characteristics they need in understanding the English aural message. EFL-LSS was designed on the basis of Perceptual Learning Preferences Survey by Kinsella (1995). Perceptual Learning Preferences Survey is a 32-item scale which was designed to help learners and their teachers understand the ways they prefer to learn. This scale comprises four learning preferences: visual/verbal, visual/nonverbal, auditory, and visual/tactile kinesthetic. Since the present study aimed to investigate the effect of different types of visuals in audio and video input on EFL students' academic listening performance, Kinsella's (1995) survey was specially chosen as a basis for the scale development because it involves visual and auditory dimensions which were essential elements of this study. Another reason of selecting Kinsella's survey as the theoretical background was its potential to produce new items into different dimensions aiming at English language listening style.

After detailed review of literature, previously developed scales focusing on determining language learners' learning styles and Kinsella's scale were examined for appropriate wording. Before developing the items in the scale, unofficial negotiations were held with colleagues and students in English language institute. Several questions were asked in order to understand how they approach to listening in English. For example; When you listen a recording in English, do you visualize them in your mind? If yes, do they help you understand better?, If a listening text involves visuals, do you think they will make the message understandable?, Is understanding every word in the recording important for better comprehension?, Which one would you prefer? A recording with visuals or a recording with no visuals (just audio). Moreover, in this itemizing process, the items in Kinsella's Perceptual Learning Preferences Survey were not translated word by word but, they were used as a reference while establishing the scale and determining the sub-dimensions of the scale. After these negotiations and studies on Kinsella's survey, a pool of 103 items aiming at determining EFL language learners' listening style were developed in accordance with Kinsella's theoretical structure. While developing the items, researcher tried to make the statements clear and easy to read by using simple language without jargons and complex expressions. After consulting with the supervisor of the study, statements which included ambiguous or leading words and irrelevant expressions were removed from the scale and the number of the items was reduced to 25.

In the process of itemizing the scale, an email which asked for an expert opinion about the content validity of these 25 items (see Appendix C) was sent to assistant professors, associate professors, and professors who were in the ELT departments of various universities. The experts were asked to evaluate items with

regard to relevance, content coverage and understandability. Based on their scrutiny and suggestions, some statements in the scale were deleted (i.e., items 1,2,3,4 in Appendix C), and reformulated and the number of the items were determined as 21 in the scale. After doing necessary revisions, the questionnaire was assigned to experts again and revised till it was considered to be satisfactory. In the scale, the participants were asked to indicate the extent to which they agree with the statements related to their listening style on a 5 point Likert scale which was set as “strongly disagree (1), disagree (2), undecided (3), agree (4), strongly agree (5)”. In terms of scoring the scale, item scores are summed for a given subscale and a mean of the summed scores is calculated for each separate subscale. Higher mean score in a subscale indicates participants’ Listening style.

3.5.2.2. Piloting and Validating the EFL Listening Style Scale

This scale was piloted twice in order to eliminate unexpected issues and test its psychometric features along with applicability. In the first piloting, the scale, which comprised 21 items, was administered to 300 EFL students who were in different English language proficiency levels (A1, A2, B1 and B2) in the same English language institute in October 2013. Since the participants seemed to have difficulties in understanding five of the items (i.e., items 17, 18, 20, 21, and 25 in Appendix C) and they did not fit into any category based on the factor analysis, these five items were deleted from the scale and three new items were added (items 6, 16, and 18 in Appendix D) after receiving an expert opinion. The validity and reliability studies were performed on the remaining 19 items (see Appendix D).

In the second piloting, 19-item scale was administered to 275 EFL students who were in A2, B1, and B2 levels in the same English language institute in May

2014. After collecting the data from the second piloting, exploratory factor analysis was carried out in order to analyze the construct validity of the scale. Before the analysis of the scale, Kaiser- Meyer-Olkin coefficient (KMO) and Bartlett's test of Sphericity tests were conducted in order to find whether data are large enough to apply a factor analysis. The KMO value varies between 0 and 1. A value being higher than .60 indicates that data are acceptable for factor analysis (Büyüköztürk, 2004; George and Mallery, 2001; Pallant, 2001). As can be seen in table 3.5, for the data gathered from the second piloting, the KMO value was found to be .729 (.729> .60) and Bartlett's test of Sphericity resulted in a significant value (.000, $p < 0.01$) which shows a high correlation between the variables. As a result, the KMO value is high (.729) and Bartlett's test is meaningful (.000), and the data are suitable for the principal components analysis.

Table 3.5 Results of KMO and Bartlett's Tests

Kaiser- Meyer-Olkin Sampling Adequacy Measure		.729
Bartlett's Test Results	X^2	778.870
	<i>Degree of Freedom</i>	136
	<i>P</i>	.000

By using principle component method and varimax rotation (rotated component matrix), explanatory factor analysis was done in order to investigate factor structure of the scale within the scope of construct validity. The result of factor analysis showed that two items (items 6 and 15) of the scale did not fit into any category which was supposed to measure a certain feature items (see Appendix D). Therefore, these two items were removed from the scale, and the number of the items in the scale was decreased from 19 to 17. Furthermore, during the analysis, item 11 was regarded as inverse item and was coded inversely. That is, the answer “strongly

agree” was graded as “1” point and the answer “strongly disagree” was 5 points. As a result, it was found that 17 items distributed into four independent factors with factor loading was greater than .32 (see table 3.6). Based on the principal components analysis results, seven items (items 1,2,3,7,9,12 and 17) were distributed to the “Visual style” and the factor loadings of these items varied between .381 and .688, indicating 15.64%. The common characteristic of these items is to identify whether EFL listeners get help from visual clues while trying to understand what they listen in English. That’s why it was labelled as “Visual Style”. In the second sub-dimension, there were four items (items 4,8, 14, and 19) which were related to the “Spatial style” and the factor loadings of these items varied between .546 and .787, explaining in 14.54% of the total variance. This construct was entitled as “Spatial Style” because it involves the items that determine whether EFL listeners use their imagination while trying to understand what they listen in English. In the third sub-dimension, three items (items 5,10, 1nd 13) were found under the “Auditory style” and the factor loadings of these items varied between .677 and .786 and also explaining in 10.88% of the total variance. The common feature of these three items is to identify how EFL listeners get help from only audio channel not visual channel while trying to understand what they listen in English. Therefore, it was called “Auditory Style”. Finally, in the fourth factor, three items (items 11, 16, and 18) were related to the “Bottom-up style” and the factor loadings of these items varied between -.522 and .796 while explaining 7.74% of the total variance. The reason why this construct was named as “Bottom-up style” is that these items identify whether EFL listeners approach the spoken texts step by step while trying to understand what they listen in English. As a result, the total amount of variance explained by these four factors was 48.82%. Moreover, the Eigenvalue graphic of the scale shows that 17 items are distributed into four sub-dimension whose eigenvalues were bigger than .33 (see

figure 3.2). This result indicates that the items in the scale generally measure similar features. (The final version of the EFL-LSS can be seen in Appendix E and F).

Table 3.6 Factor Analysis Results of the Scale-Rotated Components Matrix

Dimensions	Item No	Factor Loadings			
		Visual Style	Spatial Style	Auditory Style	Bottom-up Style
Visual Style	1	.688	-.143	-.043	.004
	2	.687	-.040	-.049	.089
	3	.650	.113	.117	-.223
	7	.630	-.080	.295	.019
	9	.504	.203	.058	.077
	12	.458	.231	.153	.046
	17	.381	.160	.114	-.087
Spatial Style	4	.078	.787	.019	.018
	8	.001	.781	.203	-.004
	14	.229	.780	-.021	-.112
	19	-.003	.546	.195	.357
Auditory Style	5	.110	.087	.786	-.020
	10	.081	.174	.717	.052
	13	.120	.006	.677	.008
Bottom-up Style	11	.020	.018	-.099	.796
	16	.329	-.230	-.146	-.522
	18	.288	-.292	.063	.439
Total variance explained (%) = 48.82		15.64	14.54	10.88	7.74

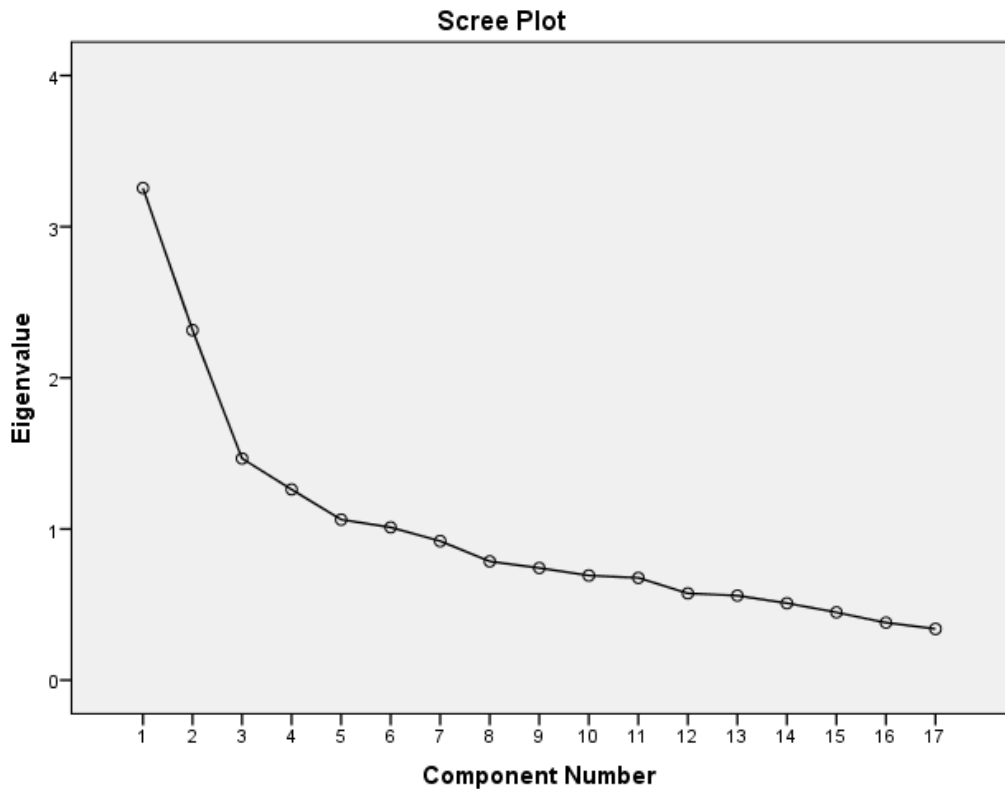


Figure 3.2. Eigenvalue graphic

3.5.2.2.1. Reliability Study of the Scale

In order to investigate the reliability of the scale, Cronbach's α reliability coefficient was calculated (see table 3.7). Based on the result of the analysis, the reliability score of the scale, which consists of 17 items, was .76. This value indicates that the scale is reliable because in the psychometric literature, it is recommended that the alpha value should be higher than .70 for the aim of research (Nunnally, 1978; as cited in Gadermann, Guhn, & Zumbo, 2012). Moreover, according to Fornell & Larcker, (1981), 0.7 is an acceptable reliability score in literature. In terms of reliability scores of sub-dimensions in the scale, Cronbach's α coefficient value was found .70 for "Visual Style", .71 for "Spatial Style", .59 for "Auditory Style", and .46 for "Bottom-up Style". The reason of why Cronbach's coefficient values of "Auditory Style" and "Bottom-up Style" are lower than the reliability score of the scale can be

justified by the fact that there is not enough number of items in these two sub-dimensions of the scale.

Table 3.7 The Reliability Values for the Scale

Sub Factors/Scales	Cronbach's α
Visual Style	.70
Spatial Style	.71
Auditory Style	.59
Bottom-up Style	.46
Scale Total	.76

3.5.3. Listening Tests

In order to investigate difference among different types of input modes (audio-only, audio with visuals and video) and also investigate difference between different types of visuals (context and content) on EFL students' academic listening performance, two listening tests which comprise 10 similar parts (2 audio only, 2 audio with content visuals, 2 audio with context visuals, 2 video with content visuals, and 2 video with context visuals) were designed for the present study.

3.5.3.1. Developing of the Listening Tests

Two listening tests which include 10 similar parts (each test involves 5 parts) and 121 multiple-choice questions in total were developed in order to investigate the impact of different types of input modes (audio and video) and different visuals (context and content), if any, on EFL students' academic listening performance. After detail review of the research studies in the literature about writing test specifications for EFL listening tests, the researcher decided to write the specifications of the present tests based on the approaches of Alderson, (1984); Hughes, (1989); and Bachman and

Palmer, (1996). According to Alderson, (1984, p.9), “A test’s specifications provide the official statement of what the test tests and how it tests it.” The reason of adapting the approaches of these researchers together was to provide more detail information about the constructs of the tests in the present study (see table 3.8).

Table 3.8 Test Specifications of the Listening Tests

Parts of the Specifications	Explanations
Purpose	To measure EFL learners’ ability to understand verbal and visual information from academic spoken texts supported by audios with context and content visuals, and videos.
Definition of the construct	The ability to understand verbal and visual information from academic spoken texts supported by audios with context and content visuals, and videos with context and content visuals in an English medium university setting.
Types of the texts	Academic mini talks
Sources of the texts	Exam sources of Q Skills for Success Level 3 Listening and Speaking, www.youtube.com , www.coursera.org
Addressees of the texts	EFL students who studied at English language institute of a private university from A2, B1, and B2 levels
Length of the texts	min.2:51- max. 3:49
Topics	Topics were selected based on the syllabus of A2, B1 and B2 level
Target level of the texts	Intermediate level
Style of the texts	Formal monologues without non-native accent, provided in American English
Characteristics of the texts and tasks	There are two listening tests, Listening Test-1 and Listening Test-2. Each test comprises 5 parts; one audio-only, one audio with context visual, one audio with content visual, one video with context visual and one video with content visual texts. Audios with context visuals comprise still images

about the context of the academic texts, whereas audio with content visuals comprise still images that semantically related to the information given in the text. Videos with context visuals involve visual information about the context of the academic texts, whereas video with content visuals involve visual information that semantically related to the information given in the video. The total number of the parts in two listening tests is 10. EFL learners will listen or watch each part twice on the screen of the Smartboard and then will choose the best answer from the three-option multiple choice questions.

Speed of processing	min. 122 w/m- max. 139 w/m
Number of items and pages	Listening Test-1: 5 parts, 50 items, 11 pages Listening Test-2: 5 parts, 45 items, 11 pages
Medium	Smartboard, paper and pencil.
Timing	Total timing is two classroom hours. Students should spend 42 minutes for the listening test-1 in the first classroom hour and 41 minutes for the listening test-2 in the second classroom hour.
Setting	Classrooms are equipped with a computer and a Smartboard
Instruction	Instruction is given by the instructor orally in Turkish at the beginning of each test. Besides, the written and oral instructions for each part of the exam will be provided in English during the test.
Scoring	Each correct response for the multiple-choice questions is assigned a value of 1, whereas each incorrect answer is assigned a value of 0.

As can be seen in figure 3.2, two similar listening tests were administered to EFL students in order to collect main data. Each test involves five parts, which name audio

only part (no visuals), audio with context visuals part, audio with content visual part, video with context visual part, and video with content visual part (see figure 3.3).

TITLE OF THE SUBTESTS	TOPIC OF THE SUBTESTS	TEST TYPE	INPUT MODE	VISUAL TYPE	LENGTH	SPEECH RATE (w/m)
“People who made an impact on a society”	Inspiring People	Test-1	Audio	Context	3:26	126
“McDonald’s”	Fast Food Industry	Test-1	Audio	Content	3:25	123
“Cigarette Smoking”	Health	Test-1	Video	Context	3:45	129
“Global Warming”	Environment	Test-1	Video	Content	3:18	122
“First Impression of Food”	Human Senses	Test-1	Audio	No Visual	3:49	124
“Greeting Cards”	Marketing	Test-2	Audio	No Visual	3:37	139
“Risk Takers”	Unusual People	Test-2	Audio	Content	3:34	130
“Climate Change”	Environment	Test-2	Video	Context	3:24	133
“Earthquake”	Disasters	Test-2	Video	Content	2:51	120
“Driving Age”	Young Drivers	Test-2	Audio	Context	3:48	128

Figure 3.3. Characteristics of the parts in the listening tests

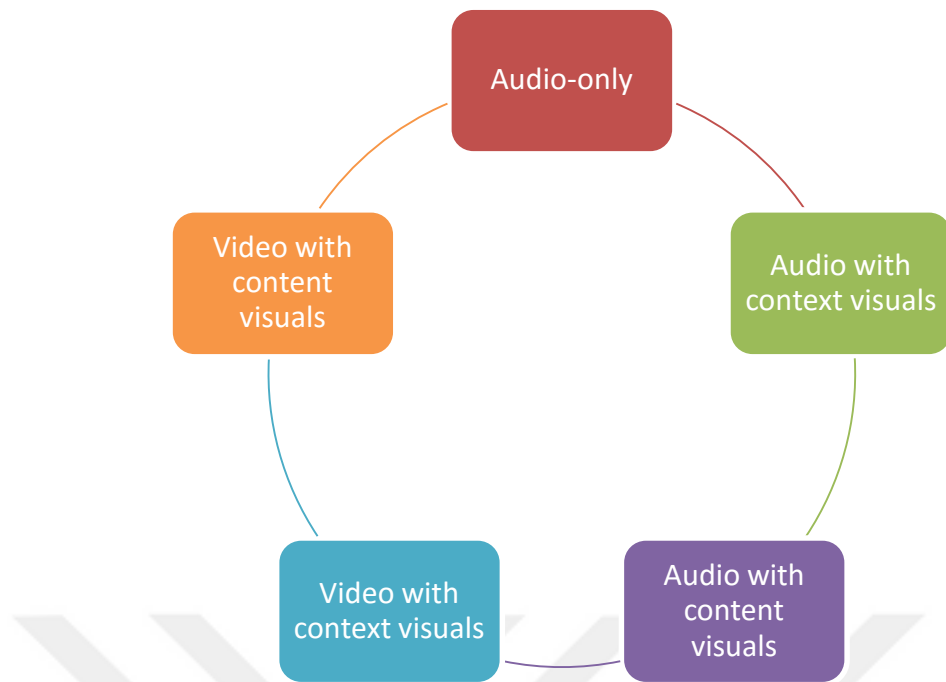


Figure 3.4. The input modes in each listening test

Many second language institutes use already published materials, which are produced by well-known publishers in ELT publishing market, in their English language instructions and also for their placement, achievement, and proficiency exams. Since ready-made listening tests provided by publishers did not meet the researcher's need for the present study, the researcher herself decided to construct a listening test which helped to achieve her aims. The listening tests, designed by the researcher herself, include both authentic and scripted materials. All audio texts (six out of ten texts) in the tests were taken from the exam materials of Q Skills for Success Level 3 Listening and Speaking from Oxford University Press, which was used as the main course book for Listening and Speaking skill instruction in B1 level in the institute, in order to minimize the potential influence of elements which are essential for successful listening comprehension, such as rate of speech, length of the recordings, accent, language level, and topic familiarity. Three videos (Global Warming, Climate Change, and Earthquake) in both tests were taken from

www.youtube.com and one video (Cigarette Smoking) was taken from www.coursera.org. The determining of the audio texts was an easy process compared to the selection of the videos. Since all the audio recordings were the exam materials of the coursebook used in the language Institute, students were familiar with the topics, structure, accent, rate of the speech, or length of the texts. Although the series of coursebook “Q Skills for Success Listening and Speaking “were used only in the B1 and B2 Levels in the institute, A2 level students were also familiar with the topics used in the books. Since the topics of the skill courses were determined based on the academic topics used in the university courses in the institute at the beginning of the academic year, students have to deal with different kinds of academic topics (e.g., environment, disasters, agriculture, economy, sports, healthcare, psychology, and management) even from the beginning of the A1 level.

On the other hand, the video recordings used in the tests were searched on the Internet and selected from hundreds of videos. The main obstacle that the researcher faced with was that although some of the videos were thematically appropriate to the learners, their contents were linguistically very difficult. Vandergrift and Goh (2012) mention that if listeners are acquainted with the form of different kinds of listening text, they can predict the structure of the discourse and use appropriate skills and strategies in order to construct meaning during listening. Therefore, the audio and video texts in the tests were taken from different sources with similar contents.

In the study, following criteria were considered when listening materials were selected and the task of the tests were designed for the test; (1) the aim of the tests, (2) the input modes, (3) types of visuals in the tests, (4) technology in the tests, (5) access to the listening texts, (6) developing test items, (7) the language level of the listening texts and the questions, (8) the proficiency level of the test takers, (9) the background

knowledge of the test takers about the listening texts in the tests, (10) any cultural effect in the texts, (11) the length of the recordings, (12) the type of listening, (13) any bias for or against specific individual, (14) the rate of the speech, (15) the audibility, (16) the accent of the speakers, (17) the number of the speaker, (18) the topic of the listening texts, (19) the familiarity of the topics, (13) the number of the questions in the test.

In the tests, academic listening texts are selected, therefore, the language in the texts is formal. Flowerdew (1994) mentions that there is a move towards informality in lectures, at least in the United States. According to the author, this trend causes problems because informality brings cultural elements, which are sometimes difficult to understand for especially non-native listeners whose cultural backgrounds are different, into the speech. Therefore, in the process of the selection of the listening materials for the tests, it was tried to be found texts free from informality and cultural elements.

When audio and video recordings were selected for the listening tests, two types of authenticity were considered. Both context and content videos in the tests are ‘adapted input authenticity’, which means that the material is created for real life but adapted by the classroom teacher. Besides, all audio recording in the tests are ‘simulated input authenticity’, which means that the material is created for the classroom and attempts to copy the style and format of the genuine (Brown, 2011).

The characteristics of the audio and video recordings and tasks in the tests are as follows:

- 1) The listening tests in this research focus on to reveal listeners’ ability to understand verbal and visual information from academic spoken texts.

- 2) In order to find whether language proficiency of the learners contribute listening comprehension or not, three different language proficiency levels- A2, B1 and B2- were included in this study.
- 3) There are 2 listening tests, each one involves 5 subtests.
- 4) Each test comprises 3 audios and 2 videos, audio only (no visuals), audio with context visuals, audio with content visual, video with context visual, and video with content visual.
- 5) The type of the questions in the tests is 'selected-response' type, which is multiple-choice question.
- 6) In the tests, there is a degree of difficulty in the items.
- 7) The places of each subtest in the tests were assigned randomly.
- 8) Context and Content photos in the audio texts were selected from the Internet.
- 9) Each audio with context and content visuals text includes series of 9-11 still images.
- 10) Few visuals in the texts contain minimum amount of textual information.
- 11) All texts are academic mini-talks varies from 2.51- 3.49 minutes.
- 12) All texts in the tests are monologue.
- 13) The researcher eliminated the effect of background knowledge by choosing familiar topics.
- 14) The difficulty of the language and vocabulary in the texts are approximately similar.
- 15) The speakers in the talks are native speaker of standard American English.
- 16) The rate of speech varies from 122- 139 w/min because Griffiths (1992) mentioned that lower-intermediate listeners performed best at 127 w/min and also Kelch (1985) found comprehension effect at 124 w/min for intermediate listeners.

- 17) Each subtest in the tests was played twice.
- 18) There are 95 questions in total.
- 19) Each test lasts approximately 42 minutes.
- 20) Each test was edited by using “Movie Maker” program and saved as mp4 files.

Decision about which audio text is going to include what kind of visuals-context or content- (or no visuals) was taken based on the topics of the audio texts. The topic which was easy to find its still pictures on the Internet was determined as audio with content visuals. For example, finding appropriate visuals about the topic “Mc Donald’s” itself is easier than finding relevant visuals about the topic “Driving Age”. Therefore, audio text about “Mc Donald’s” and “Risk Takers” were assigned to audio with content visuals (see figure 3.5 and figure 3.6), the texts about “People who made an impact on a society” and “Driving Age” were assigned to audio with context visuals (see figure 3.7. and figure 3.8), and the topics “First impression of food” and “Greeting Cards” were named as audio only (no visuals) texts.

Furthermore, decision about which audio and video texts were context or content were made based on the definition of context and content visuals presented by Bejar et. al. (2000) and Ginther (2002). Audio and video texts displayed a lecturer and the places where the lecture was given (“Driving Age” and “People who made an impact on a society” for audios and “Cigarette Smoking” and “Climate Change” for videos) were named as “audio with context visuals” (see figure 3.7 and figure 3.8) and “video with context visuals” (see figure 3.9 and figure 3.10). Moreover, audio and video texts displayed visuals related to the topics of the texts (“Mc Donald’s” and “Risk Takers” for audios and “Global Warming” and “Earthquake” for videos) were selected as “audio with content visuals” (see figure 3.5 and figure 3.6) and “video with content visuals” (see figure 3.11 and figure 3.12).



Figure 3.5. A screenshot from audio with content visuals in Test-1



Figure 3.6. A screenshot from audio with content visuals in Test-2



Figure 3.7. A screenshot from audio with context visuals in Test-1



Figure 3.8. A screenshot from audio with context visuals in Test-2



Figure 3.9. A screenshot from video with context visuals in Test-1



Figure 3.10. A screenshot from video with context visuals in Test-2

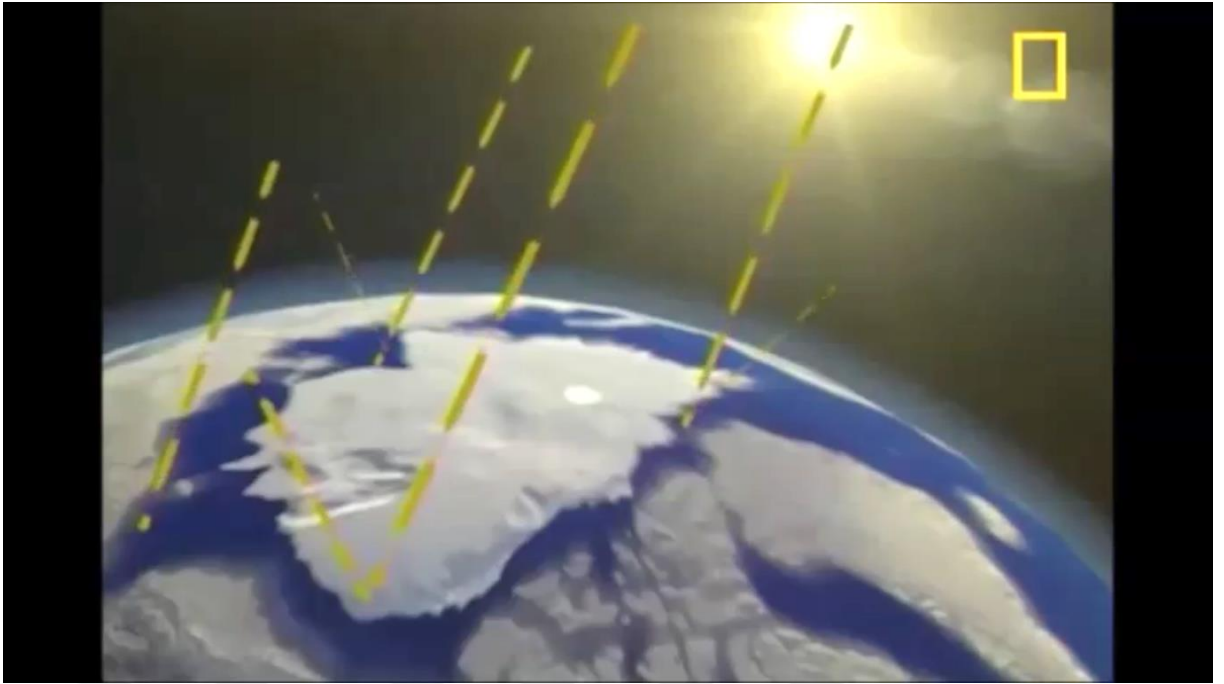


Figure 3.11. A screenshot from video with content visuals in Test-1



Figure 3.12. A screenshot from video with content visuals in Test-2

In order to assign the difficulty level of the listening texts-audios and videos-used in the present study was measured by Flesch Reading Ease Score and Flesch-Kincaid Grade Level formulas. These formulas are widely used readability formulas in order to determine the difficulty of a written text. Although these formulas were designed to assess the readability of a written text, they are used to assess the listenability by many researchers in the field (Suvorov, 2013). Rubin (1993) also states that although there are some worries about inadequacy of “empirical validations of listenability/readability equation”, these formulas are still applied into research. According to DuBay (2007) (as cited in Suvorov, 2013), even though listenability and readability were found to be correlated by many researchers, these formulas are measured readability not listenability. Therefore, the obtained scores must be interpreted cautiously in terms of listenability.

In the Flesch Reading Ease formula, scores are scaled from 0 to 100. Higher scores mean that reading texts are easy to read, on the other hand, lower scores indicate that the texts are difficult in terms of readability. Moreover, the Flesch-Kincaid Grade level formula ranked the scores of the reading texts as a US grade levels. Higher grade levels mean that texts are difficult to read.

The Flesch Reading Ease Scores and the Flesch-Kincaid Grade Levels of the audio scripts of audios and videos in the listening tests are given in Table 3.9 The descriptions of the scores are taken from Flesch (1948). According to the Ease scores and grade level of the listening text, video with content visuals in the Listening Test-1 is difficult and audio with context visuals in Listening test-1 is fairly difficult. Moreover, video with context visuals in Listening Test-2 is fairly difficult. Since all EFL students from all proficiency levels took all these subtests, this was not a cause for concern.

Table 3.9 Readability Statistics for the Scripts of Audios and Video Clips in the Tests

Test	Input Type	Visual Type	Word Count	Flesch Reading Ease Score	Flesch-Kincaid Grade Level	Reading Score Explanation
Test-1	Audio	Context	416	54.1	9.1	Fairly Difficult
Test-1	Audio	Content	436	59.7	8.1	Standard
Test-1	Video	Context	526	65.5	5.0	Standard
Test-1	Video	Content	412	48,2	10.3	Difficult
Test-1	Audio	No visual	494	76.9	5.8	Fairly Easy
Test-2	Audio	No visual	615	68.4	7.0	Standard
Test-2	Audio	Content	493	69.6	6.9	Standard
Test-2	Video	Context	463	52.7	12.5	Fairly Difficult
Test-2	Video	Content	338	65.7	7.5	Standard
Test-2	Audio	Context	531	61.8	8.3	Standard

Each of two listening test comprises 5 subtests (audio-only, audio with context visuals, audio with content visuals, video with context visual, and video with content visual). Since two identical listening tests were administered, there were 10 subtests (i.e., 2 audio-only, 2 audio with context visuals, 2 audio with content visuals, 2 video with context visual, and 2 video with content visual) in the tests, as can be seen in figure 3.13 and figure 3.14. At the very beginning of the designing process of the listening tests, only one listening test, which included 5 subtests (audio-only, audio with context visuals, audio with content visuals, video with context visual, and video with content visual), was going to be used as a data collection instrument of this study.

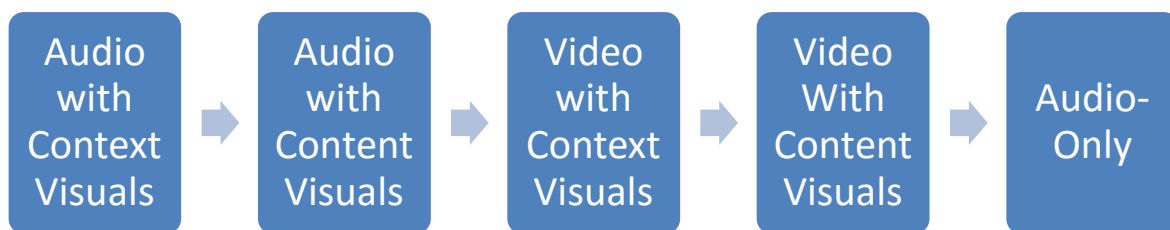


Figure 3.13. The order of the subtests in the Listening Test-1

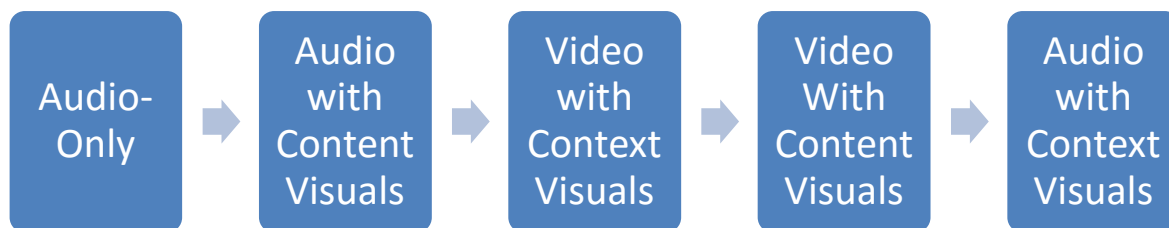


Figure 3.14. The order of the subtests in the Listening Test-2

The reason for constructing two identical listening tests, which involved similar subtests in terms of level of difficulty, the duration, or the rate of speech, was to increase the reliability of the tests. In other words, Associated Professor Hossein Farhady, who is in the field of testing in English Language Teaching Department, mentioned that the number of the questions in each subtest of a single test was not enough to provide a reliable test and a test with more test items will increase its reliability. Therefore, he advised increasing the number of the test items to at least 15 to 20 for each subtest. However, the biggest obstacle that we faced with in the construction process of the test was that the length of the single audio or video was not long enough to produce 15 to 20 test items for each subtest. Therefore, the Associated Professor suggested that designing one more listening test which includes similar recordings as the first test comprises would be the best option to increase the number of the items in the tests. As Tavakol and Dennick (2011) mention that adding more related items which are testing the same concept to the test will increase the reliability score of the test. Furthermore, Taylor and Geranpayeh (2011) indicate that rather than providing one long listening text, constructing the academic listening test

with a series of shorter listening texts should be an ideal occasion for test takers. As a result, two identical listening tests, which comprise 10 similar subtests, were started to be constructed.

All ten subtests of two listening tests include multiple choice questions. One of the reasons why multiple-choice task was used in the test was to minimize the effect of other skills such as speaking and writing on the tests (Dunkel, Henning, and Chaudron, 1993). Secondly, since the duration of all two listening tests is long (approximately two class hours), answering to multiple-choice questions must not take time. Finally, since the number of the students who took the exam was high, multiple choice tasks eased the scoring process. Each multiple choice question in the test comprises a stem in a form of either an incomplete statement (Smoking was first popular among _____.) or a direct question (Why do some people refuse to eat sushi?). All questions in the tests involve three options with one correct response (key) and two distractors. The questions in each listening text were developed to elicit both the main idea (e.g., What is the main idea of the lecture/talk?) and details of the talks (e.g., How many customers visit McDonald's every day? or How much has the Arctic Sea ice decreased according to NASA studies?)

As it was mentioned in the above sections, during the material selection process, several resources had been investigated in terms of finding appropriate audio and video materials for the listening tests. After listening and watching hundreds of audios and videos, the researcher decided to select all audio recordings of the listening tests from the test material sources of the coursebook "Q Skills for Success Level 3 Listening and Speaking", three video recordings from www.youtube.com and one video from www.coursera.org. While selecting the audio and video materials, the researcher was very careful about the language levels of them. Since the audio

recordings were the exam materials of the coursebook, which was already used in B1 and B2 levels, and also the topics in these recordings covered the content of A2, B1, and B2 levels, the researcher and her colleagues, who were the members of the Testing Office in the university, found the audio recordings appropriate in terms of content, the level of difficulty, rate of speech, background knowledge, and accent, for all the levels which participated in the study. Once audio/video recordings and accompanying visuals were selected, all these materials were transformed to Mp4 file and edited by using “Movie Maker” program.

After saving the recordings and visuals as Mp4 files, the researcher sent them to three senior examiners who were also experienced EFL instructors (one of them was a native speaker of English language) in the Testing Office of the university in order to increase the number of the multiple-choice questions for each recording. Since there was not enough number of questions for the audio recordings, which were the listening test materials of the coursebook “Q Skills for Success Level 3 Listening and Speaking”, and there were not any questions for the video recordings, the researcher herself and the members of the Testing Office began to produce new test items for both audio and video recordings. Furthermore, Alderson, Clapham and Wall (1995, p. 69) mentioned “Nobody writes good test items alone” Therefore, people who are experts on their field were involved in constructing, producing, checking, editing, proof reading, developing, designing, and testing process of the listening tests. During the process of writing the new test items, the test writers tried to make the questions and options clear, understandable, and easy to read by using simple language without complex expressions. After completing the first draft of the listening tests, the committee came together in order to ensure the test items with their options in the tests were fair, clear, understandable, and error free. The committee also checked

whether the still pictures in the audios and the images in the videos were appropriate for the topics and determined whether the visuals were accordance with the spoken messages. The committee assessed the tests in terms of language level and task difficulty as well. After finalizing the construction of the first draft of the tests with their answer keys, they were sent to an expert in the field of Testing to get his opinions and suggestions about what the testing committee designed. The expert worked on the draft and he edited and moderated the items and also made suggestions for changes where necessary. Then, the vetted copy was sent to the testing committee again for proof reading and for final check based on the expert's report. Once the agreed copy of the listening tests was completed, the order of each audio and video subtest in the tests was randomly assigned. Later, the first version of the tests was typed on Microsoft Office Word 2016 and printed multiple copies for piloting (see Appendix Q and R).

3.5.3.2. Piloting the Listening Tests

Sixty-one EFL students were recruited for the pilot study of the listening tests. The participants in the piloting were in different language levels of English (A2, B1 and B2). The overall profiles of the students in the pilot study were similar to the profiles of the students who participated in the main study. The pilot data was collected between May 26th-29th, 2014 in the same private university.

The initial version of the listening tests involved 142 questions in total (69 questions in the first test, 73 questions in the second test). After getting necessary permission from the director of the language institute, the piloting process started. First of all, the date and the time line for the administration of the tests were assigned. Then, the classes from three different language levels were determined randomly (one

A2, two B1, and two B2 levels). Later, instructors who had lessons on the selected classes were trained about the administration and the procedure of the listening tests. Finally, Mp4 files of the listening tests were uploaded into the computers in every classroom, and the test booklets were printed.

At the beginning of the administration of the tests, the booklets of the first listening test were distributed to the participants and all instructions were given in Turkish by the classroom instructors. During the testing process, all subtests in the test were played twice. The participants were allowed to see the questions and to answer them on the given booklet while the audio and video recordings were playing during the tests. All two tests, listening test-1 and 2, were carried out in two classroom hours. Since each test took approximately 42 minutes, 50-minute classroom hour was enough to complete each test. After the administration of the listening test-1 in the first classroom hour, 20-minute tea break was given. After the break, the second listening test booklets were delivered and then the administration of the listening test-2 was carried out. After completing all subtests of the listening test-1 and 2, the booklets were collected by the instructors and were handed in to the researcher herself. Later, all the papers were double checked by two assessors (each correct answer was scored as “1” point whereas each incorrect response was scored as “0” point) and test scores of the pilot study were transferred to Microsoft Office Excel 2016 and SPSS 16.0 programs in order to analyze them.

3.5.3.2.1. Data Analysis of the Piloted Listening Tests

The analyses of the listening tests data were done by examining the item difficulty, item discrimination, distractor distribution of the multiple choice items, and reliability scores of each part in two listening tests. The content validity of each part

of two listening tests was carried out as well. In total, three kinds of statistical analyses (i.e., item analysis, distractor analysis, and reliability analysis) were conducted by using the performance data from the listening test-1 and test-2 (N=61).

Item analysis. In order to investigate the properties of the items in the listening test-1 and the listening test-2, item analyses were implemented using the item scores on both piloted tests (for the piloted items on both tests see Appendix Q and Appendix R). As can be seen on table 3.10 and table 3.11., item facility (IF), which is also called as item difficulty, and item discrimination, which was calculated by using point-biserial correlational coefficient (r_{p-bis}), for each item on both tests are presented in order to develop their final format.

Table 3.10 Item Facility and Item Discrimination for Each Item on the Listening Test-1

Listening Test-1			
Parts of the Test-1	Item No	IF	ID
Part-1 Context Audio	1	0.738	0.625
	2	0.800	0.401
	3	0.969*	0.162
	4	0.969*	0.014
	5	0.523	0.558
	6	0.815	0.485
	7	0.615	0.492
	8	0.769	0.600
	9	0.785	0.339
	10	0.738	0.493
	11	0.769	0.341
	12	0.954*	0.155
	13	0.169*	0.349
	14	0.431*	0.099
	15	0.646	0.469
Part-2 Content Audio	1	0.808	0.621
	2	0.699	0.412
	3	0.973*	-0.058
	4	0.808	0.455
	5	0.795	0.646
	6	0.603	0.496
	7	0.603	0.521

	8	0.753	0.455
	9	0.630	0.312
	10	0.849*	0.214
	11	0.740	0.320
	12	0.877*	0.443
	13	0.575	0.388
Part-3 Context Video	1	0.756	0.707
	2	0.837*	0.734
	3	0.837*	0.657
	4	0.767	0.792
	5	0.744	0.667
	6	0.291*	0.389
	7	0.779	0.283
	8	0.756	0.460
	9	0.709	0.757
	10	0.756	0.468
	11	0.709	0.446
	12	0.698	0.543
	13	0.942*	-0.096
	14	0.744	0.772
Part-4 Content Video	1	0.246*	0.118
	2	0.918*	0.330
	3	0.787	0.564
	4	0.689	0.576
	5	0.770	0.547
	6	0.279*	0.310
	7	0.770	0.649
	8	0.852*	0.469
	9	0.754	0.419
	10	0.738	0.554
	11	0.787	0.459
	12	0.590	0.647
	13	0.492	0.515
	14	0.590	0.440
	15	0.426*	0.168
Part-5 Audio-Only	1	0.875*	0.128
	2	0.750	0.556
	3	0.734	0.499
	4	0.781	0.591
	5	0.703	0.461
	6	0.453	0.643
	7	0.203*	0.135
	8	0.563	0.582
	9	0.781	0.349
	10	0.516	0.554
	11	0.766	0.351
	12	0.594	0.544

Note. Deleted test items are in bold and marked with an asterisk (*)

Table 3.11 Item Facility and Item Discrimination for Each Item on the Listening Test-2

Listening Test-2			
Parts of the Test-2	Item No	IF	ID
Part-1 Audio-Only	1	0.982*	-0.005
	2	0.946*	0.095
	3	0.857*	0.185
	4	0.732	0.418
	5	0.839*	0.175
	6	0.589	0.397
	7	0.429	0.401
	8	0.732	0.524
	9	0.625	0.341
	10	0.875*	0.198
	11	0.696	0.604
	12	0.750	0.501
	13	0.625	0.551
	14	0.357	0.421
Part-2 Content Audio	1	0.913*	0.697
	2	0.797	0.326
	3	0.971*	0.314
	4	0.551	0.635
	5	0.783	0.412
	6	0.870*	0.502
	7	0.841*	0.432
	8	0.913*	0.712
	9	0.551	0.592
	10	0.739	0.713
	11	0.841*	0.455
	12	0.870*	0.578
	13	0.739	0.751
	14	0.768	0.456
	15	0.565	0.558
	16	0.478	0.394
Part-3 Context Video	1	0.966*	0.248
	2	0.190*	0.296
	3	0.672	0.531
	4	0.690	0.593
	5	0.155*	0.018
	6	0.552	0.575
	7	0.741	0.413
	8	0.724	0.381
	9	0.724	0.480

	10	0.759	0.568
	11	0.655	0.626
	12	0.741	0.497
Part-4 Content Video	1	0.903*	0.439
	2	0.226*	0.007
	3	0.871*	0.148
	4	0.500	0.382
	5	0.774	0.593
	6	0.742	0.332
	7	0.500	0.712
	8	0.645	0.670
	9	0.871*	0.352
	10	0.677	0.552
	11	0.645	0.479
	12	0.968*	0.288
	13	0.903*	0.227
	14	0.903*	0.304
	15	0.742	0.735
	16	0.597	0.438
Part-5 Context Audio	1	0.867*	0.291
	2	0.633	0.662
	3	0.950*	0.099
	4	0.967*	0.347
	5	0.800	0.379
	6	0.850*	0.126
	7	0.617	0.626
	8	0.650	0.603
	9	0.750	0.580
	10	0.850*	0.455
	11	0.817	0.480
	12	0.683	0.555
	13	0.783	0.427
	14	0.900*	0.057
	15	0.650	0.357

Note. Deleted test items are in bold and marked with an asterisk (*)

According to Bachman (2004), item facility (IF) values should fall in a range between .20 and .80 and item discrimination indices (ID) should be greater than .30. Based on these given scores, the piloted data of both listening test-1 and listening test-2 were analyzed to select the best items and to design more effective version of the tests. As can be seen in table 3.10 and table 3.11, the items that fell within a range between .20 and .80 in item facility and had the highest item discrimination index were kept for the new version of the tests. As can be seen in table 3.10 and table 3.11,

the items written in bold rejected because of low item discrimination index and very low or very high item facility value. These rejected items deleted completely for the new and the final version of the tests. Yet, only four items with a value over .80 in two tests were retained for future use because McNamara (2000) states that keeping some very easy items, especially at the beginning of the test, may be useful for students to give them a chance to handle their stress. Therefore, the item 6 in Part-1 Context Audio and the items 1 and 4 in Part-2 Content Audio in Listening Test-1 (see table 3.10) and the item 11 in Part-5 Context Audio in Listening Test-2 (see table 3.11) were kept for the final version of the tests. Meanwhile, the lowest item facility value which will be placed in the final version of the tests was .35 (item 14 in Part-1 Audio-Only in Listening Test-2) as shown in table 3.11. It means that very difficult questions were not placed in both future tests.

Regarding the item discrimination, only one item, which will be in the final version of the tests, was found to have low point-biserial correlation coefficients (see item 7 in Part-3 Context Video in Listening Test-1). Ebel (1979) (as cited in Brown, 2005) defines items which fall a range between .20 and .29 as “marginal item”. Hence, the item which had low ID value (.28) in Part-3 Context Video in Listening Test-1 may not discriminate lower-level ability EFL students from the students who have higher-level ability in EFL listening. Yet, this item was kept for the final version of the tests because of its IF value (.779). The number of the questions in each subtest and the total number of the questions in both listening test-1 and test-2 after the item analysis are presented in table 3.12. As can be seen in the table 3.12, the number of the questions in each subtest of the listening test-1 and the listening test-2 are combined in the future analyses for the main study.

Table 3.12 Total Number of the Questions in each Subtest of both Listening Tests After the Item Analysis.

Subtests of the Listening Tests					
	Subtest-1	Subtest -2	Subtest -3	Subtest -4	Subtest -5
Listening Test-1	Audio with	Audio with	Video with	Video with	Audio-Only
	Context Visuals	Content Visuals	Context Visuals	Content Visuals	
	10 Questions	10 Questions	10 Questions	10 Questions	10 Questions
Listening Test-2	Subtest -5	Subtest -2	Subtest -3	Subtest -4	Subtest-1
	Audio with	Audio with	Video with	Video with	Audio-Only
	Context Visuals	Content Visuals	Context Visuals	Content Visuals	
	9 Questions	9 Questions	9 Questions	9 Questions	9 Questions
Total Number of the Questions in both Tests	Total Context Audio 19	Total Content Audio 19	Total Context Video 19	Total Context Video 19	Total Audio-Only 19

To sum up, the item analysis of the listening test-1 and the listening test-2 showed that the difficulties of items and the capability of them (after the rejected items deleted from the tests) to separate test takers with low-level EFL listening skill from the ones with high-level EFL listening ability were appropriate for contributing to more reliable tests.

Distractor analysis. After deleting the problematic items in terms of item difficulty and item discrimination from listening test-1 and listening test-2, distractor analysis was employed for the distractors of 50 items in the listening test-1 and of 45 items in the listening test-2. This analysis was carried out to reveal any ambiguous distractors and keys (correct option). As shown in table 3.13 and table 3.14, the frequency distributions were presented for each of 50 items in listening test-1, which involved 150 options (50 correct answers and 100 distractors) and for each of 45

items in listening test-2, which included 135 options (45 correct answers and 90 distractors). According to Haladyna and Downing (1993) and Rodriguez (2005), employing 3-option multiple choice tests improve the quality of distractors because writing more plausible distractors in 3-option test is easier than writing plausible distractors in 4 or 5 option test and this type of tests recommended by several researchers (e.g., Bruno & Dirkzwager, 1995; Farhady & Shakery, 2000; Haladyna and Downing, 1993; Lord, 1977). Therefore, in these listening tests, 3-option multiple choice items were constructed.

Table 3.13 Results of Distractor Analysis of Item Scores on the Listening Test-1

		Listening Test-1		
Parts of Test-1	Item No	Response Frequencies		
		%A	%B	%C
Part-1 Context Audio	1	15.38	10.77	73.85
	2	80.00	12.31	7.69
	3	21.54	52.31	26.15
	4	12.31	6.15	81.54
	5	61.54	24.62	13.85
	6	9.23	13.85	76.92
	7	78.46	10.77	10.77
	8	16.92	73.85	9.23
	9	76.92	13.85	9.23
	10	64.62	20.00	15.38
Part-2 Content Audio	1	6.85	12.33	80.82
	2	5.48	24.66	69.86
	3	9.59	80.82	9.59
	4	79.45	6.85	13.70
	5	60.27	9.59	30.14
	6	23.29	60.27	16.44
	7	13.70	75.34	10.96
	8	21.92	63.01	15.07
	9	17.81	8.22	73.97
	10	57.53	19.18	23.29
Part-3 Context Video	1	16.28	75.58	8.14
	2	10.47	76.74	12.79
	3	74.42	13.95	11.63
	4	11.63	10.47	77.91
	5	10.47	13.95	75.58
	6	70.93	12.79	16.28
	7	13.95	75.58	10.47
	8	70.93	17.44	11.63
	9	69.77	17.44	12.79

	10	74.42	11.63	13.95
Part-4 Content Video	1	13.11	78.69	8.20
	2	68.85	8.20	22.95
	3	77.05	13.11	9.84
	4	8.20	14.75	77.05
	5	9.84	14.75	75.41
	6	13.11	13.11	73.77
	7	78.69	9.84	11.48
	8	14.75	26.23	59.02
	9	26.23	49.18	24.59
	10	18.03	22.95	59.02
Part-5 Audio-Only	1	75.00	17.19	7.81
	2	14.06	73.44	12.50
	3	78.13	7.81	14.06
	4	12.50	17.19	70.31
	5	32.81	45.31	21.88
	6	15.63	28.13	56.25
	7	78.13	10.94	10.94
	8	20.31	51.56	28.13
	9	14.06	76.56	9.38
	10	59.38	26.56	14.06

Note. Keys are in bold

Table 3.14 Results of Distractor Analysis of Item Scores on the Listening Test-2

Listening Test-2				
Parts of Test-2	Item No	Response Frequencies		
		%A	%B	%C
Part-1 Audio-Only	1	73.21	16.07	10.71
	2	58.93	26.79	14.29
	3	25.00	32.14	42.86
	4	73.21	14.29	12.50
	5	21.43	62.50	16.07
	6	14.29	16.07	69.64
	7	75.00	10.71	14.29
	8	62.50	16.07	21.43
	9	30.36	33.93	35.71
Part-2 Content Audio	1	10.14	79.71	10.14
	2	10.14	34.78	55.07
	3	15.94	5.80	78.26
	4	55.07	27.54	17.39
	5	14.49	73.91	11.59
	6	10.14	73.91	15.94
	7	10.14	13.04	76.81
	8	20.29	56.52	23.19
	9	33.33	18.84	47.83

Part-3 Context Video	1	17.24	67.24	15.52
	2	68.97	18.97	12.07
	3	55.17	27.59	17.24
	4	13.79	74.14	12.07
	5	72.41	13.79	13.79
	6	13.79	72.41	13.79
	7	75.86	10.34	13.79
	8	17.24	17.24	65.52
	9	15.52	74.14	10.34
Part-4 Content Video	1	50.00	32.26	17.74
	2	77.42	8.06	14.52
	3	16.13	74.19	9.68
	4	29.03	20.97	50.00
	5	16.13	19.35	64.52
	6	25.81	6.45	67.74
	7	15.00	65.00	20.00
	8	17.74	74.19	8.06
	9	24.19	16.13	59.68
Part-5 Context Audio	1	26.66	10.00	63.33
	2	79.66	10.17	10.17
	3	21.67	61.67	16.67
	4	15.00	20.00	65.00
	5	75.00	10.00	15.00
	6	81.67	10.00	8.33
	7	20.00	11.67	68.33
	8	10.00	78.33	11.67
	9	13.33	21.67	65.00

Note. Keys are in bold.

With regard to the distractor analysis, the proportion of all keys (correct options written in bold in table 3.13 and table 3.14) were equivalent to IF values of the items (are shown in table 3.11 and table 3.12). Regarding the distractors in the tests, different types of methods for evaluating non-functioning distractors are used in the literature. In this study, response frequency was employed to determine the quality of the distractors. According to Haladyna and Downing, (1993) and Tarrant, Ware and Mohammed (2009), non-functioning distractors have low response frequencies and if the frequencies of these distractors are less than %5, they are probably implausible ones. As can be seen in table 3.13 and table 3.14 there are no distractors which have less than %5 response frequency in both tests. Therefore, all the distractors in both

tests were plausible, functioning effectively, and were retained for the final version of the tests.

Reliability analysis. In order to investigate the reliability of the tests, Cronbach's α reliability coefficient and standard error of measurement (SEM) for each part of listening test-1 and listening test-2, as well as for the overall test scores were calculated. Based on the result of the analyses, the reliability score for the listening test-1, which consists of 50 items, was .84 and standard error of measurement 2.81 and for the listening test-2, which consists of 45 items, was .83 and standard error of measurement 2.87 (see in table 3.15).

Table 3.15 Results of Reliability Analysis of the Overall Scores and Scores of Each Part on the Listening Test-1 and Listening Test-2.

	Types of Score	Number of Test Items	Mean	Cronbach's Alpha (α)	Standard Error of Measurement (SEM)
Listening Test-1	Part-1 Context Audio	10	7.20	.68	1.27
	Part-2 Content Audio	10	6.43	.62	1.23
	Part-3 Context Video	10	6.67	.81	1.02
	Part-4 Content Video	10	6.96	.77	1.24
	Part-5 Audio Only	10	6.64	.72	1.30
Listening Test-2	Part-1 Audio Only	9	5.53	.60	1.28
	Part-2 Content Audio	9	5.97	.74	1.20
	Part-3 Context Video	9	6.25	.70	1.22
	Part-4 Content Video	9	5.82	.74	1.23
	Part-5 Context Audio	9	6.38	.71	1.19
	Overall Listening Test-1	50	35.11	.84	2.82
	Overall Listening Test-2	45	31.32	.83	2.87

In the literature, different information for the acceptable value of alpha, generally between .70 and .95, has been mentioned (Sueyoshi & Hardison, 2005; Tavakol & Dennick, 2011). However, Lado (1961) stated that acceptable reliability score for auditory comprehension test is in the .80 to .89 range. Considering the reliability scores of each part in the listening test-1 and listening test-2, as shown in table 3.15, it can be concluded that Cronbach's α reliability coefficient values of 7 out of 10 parts of the tests- except Part-1 Context Audio (.68) and Part-2 Content Audio (.62) in the listening test-1, and Part-1 Audio Only (.60) in the listening test-2- were in acceptable range for a comprehension test. This finding demonstrates that although the reliability scores on these three tests were slightly lower than the scores on other 7 parts, the overall scores of the listening test-1 (.84) and listening test-2 (.83) had adequate values in terms of reliability. Furthermore, as can be shown in table 3.15, standard error of measurement (SEM) of all parts of the listening tests and of overall scores of listening test-1 and listening test-2 are very low. According to Brown (2005), the value of SEM is more important than any reliability scores for real-life decisions because smaller SEM represents the students' actual abilities more accurately and consistently. Thus, it can be concluded that these two listening tests are reliable and measure the students' EFL listening ability accurately based on the statistical analysis of the piloted data.

Validity of the tests. At the beginning of the academic year of 2013-2014, the curriculum of the language institute and the syllabus of each language level (i.e., A1, A2, B1, and B2) for each module were ready. In order to decide the contents and the topics of each part in the listening tests and construct the test items based on the curriculum and syllabi, a meeting was requested from the administration of the language institute. At the beginning of September 2013, a meeting was held with level

coordinators for specifying the common topics for A2, B1 and B2 levels. In the meeting, all the contents and the topics of the courses, which are going to be taught through the academic year, were asked to coordinators in order to reveal the common topics among these three language levels. Later, all common academic contents and the topics among A2, B1 and B2 levels were identified and in which dates these topics were going to be taught for each level were clarified with the level coordinators. The teaching of these shared topics in A2, B1, and B2 levels until the data collection process of the main data was very important because all the audio and video texts and their test items were constructed based on these predetermined contents and topics. Otherwise, the participants of the study would take a listening test which was not valid in terms of content. Consequently, all identified academic topics of both audio and video texts were taught in A2, B1 and B2 levels before the starting of the administration of the listening tests.

In the language institute where this research study was conducted, teaching of academic language starts at A2 level and continues towards at the end of B2 level. Therefore, the academic topics, such as global warming, climate change, healthcare, economy, poverty, or famine and their related words are studied in every language level with varying levels of difficulty. Thus, all the topics of the recordings, either in the audio or video mode, in the listening tests were determined based on the syllabi of A2, B1, and B2 levels. As it was explained in table 3.8, the test specifications of the listening tests, all of six audio recordings were taken from the exam sources of the book “Q Skills for Success Listening and Speaking” which was studied in B1 and B2 levels. Thus, it can be stated that the topics and the contents of the audio recordings in the tests are completely align with the content in the syllabi of the levels.

Furthermore, although all four videos were found from the Internet, their topics and

contents were in parallel to syllabi as well. Since the contents of the courses, which were offered in A2, B1, and B2 levels, were represented in the listening tests, and the items of these listening tests were relevant to the content of these courses as well, it is possible to mention that the listening test-1 and the listening test-2 measured what we intended to measure, in terms of content validity in this research study.

3.5.4. Interview

In order to investigate the EFL students' experiences about the presence and the absence of context and content visuals provided by multi-modal input in the listening tests, semi-structured group interviews were carried out with twenty-four EFL students from 6 classrooms (2 randomly selected classrooms from each level) in Turkish. Field (2008) states that listening teachers might ask 'why' questions to the learners in order to reveal their interpretation of recordings after they answer multiple choice questions. Therefore, just after completing the listening tests, the students were asked to participate in an interview in order to talk about their experiences about the tests and their administrations. In the interview session, the EFL students who had just completed the listening test-1 and the listening test-2 were asked questions about the administration of the listening tests (e.g., What do you think about the listening tests such as the length of it, playing the recordings twice, the topics of the texts, the accent and the speech rate of the recordings?), visuals in the tests (e.g., What do you think about the visuals in the tests? Do you think visuals, which are about the topics of the recordings, are related to the topics of the texts? Did visuals help you understand the texts better or less?), the presence and the absence of the visuals in the listening tests (e.g., Some of the recording did not involve visuals but most of them had, how did the presence and the absence of the visuals affect you and your understanding?), the participants' preferences in terms of visuals (e.g., Which kind of visuals helped you

more or less, context or content? What kind of visuals do you prefer in listening tests, the one with content visuals or the one with context visuals? Would you like to listen audio recordings with visuals or without visuals? Which one would you prefer in listening tests (videos or audios?), the quality of the audios and videos (e.g., What do you think about the quality of the recordings? Did something in the audios or videos prevent you from understanding?), the topic of the recordings (e.g., What do you think about the topics of the recordings? Was there anything that you were not familiar in the recordings in terms of the topics and contents of them?), and the difficulty of the test items (e.g., What do you think about the level of the questions, were they easy/normal/difficult?). The language in the interview sessions was Turkish. A copy of the semi-structured interview questions is provided in Appendix U.

3.6. Data Collection Procedures

After getting the approval from the administration of the language institute, all four sets of data (i.e., Demographic information data from Demographic Information Questionnaire; the EFL students' listening style data from the EFL Listening Style Scale; test performance data from the Listening Test-1 and Test-2; and finally, verbal data from semi-structured interviews) were collected between January 5th, 2015 and March 27th 2015. For the study, 3 classes for each level (i.e., 3 classes for A2 level, 3 classes for B1 level and 3 classes for B2 level) were assigned. Although the classes in A2 and B1 language levels were randomly selected, all of three B2 level classes were included into the study. Before starting to collect data, timeline (including the date and the classroom hours) was planned for each class. All data except interviews were collected in the second and the third classroom hour in order to reach maximum number of participants in each class. Since there may be late comers in the first hours

and early leavers in the last hours, starting the data collection in the second hour seemed more feasible. Since all the data were collected from 9 classes by the researcher herself, arrangement of the date and the time of the administration were done with classroom instructors. Therefore, a meeting was held with the classroom instructors. In this meeting, these instructors were informed about the administration procedure of the data collection and they were asked to the best date for the administration. After specifying the time line of the policy, the first implementation started in January 5th, 2015.

Since all students were already informed about the study and the implementation process of it by their classroom instructors before, the researcher did not explain the aim of the study or the reason of the data collection procedure in detail. Yet, a brief information about how the implementation was going to be done and what kind of listening test it was were provided to the EFL students and questions, if there was any, were answered clearly in Turkish.

The data collection started with the distribution of the first booklet, which involved Demographic Information Questionnaire and EFL Listening Style Scale, and the second booklet, which was the listening test-1. All students finished to fill in the first booklet approximately in 12 minutes. Then, six minutes was given for investigating the second booklet (i.e., the Listening Test-1) and then a brief information was given about the listening tests. Later, mp4 files which had already been downloaded into the computers in the classroom started to play. All visual images and videos were played on the screen of the Smartbord and the sound propagation was provided through a surround system. All information about each part of the listening test (e.g., what the topic of the recording is, how many times it is played, how much time is needed to scan the questions in the booklet before the

listening starts, when will students start the next part, or what kind of recording it is (audio-only, audio with visuals, or video) was provided by the played file to the students. Therefore, the students were informed that any question coming from them about the listening tests during the administration was not going to be answered by the researcher. During the test phase, the participants were allowed to take notes and to use them to answer the questions.

As it was mentioned in the literature chapter, world knowledge or topic knowledge help listener build meaning. According to Field (2008), this knowledge can be stimulated before listening to text by asking students to guess what the speaker will talk about. This activity helps listeners identify possible script and think about words related to topic. If the recording is non-participatory type of listening, a lecture or a film, a competent listener might activate her background knowledge for the topic. Vandergrift and Goh (2012) also mention that giving information about the listening text at the beginning of the playing activate listeners' background knowledge and this attempt facilitate top-down processing. Therefore, in the listening tests in this study, a brief information about what students were supposed to listen was given by the played mp4 file at the beginning of each part of the test. Thus, the participants could think about the topic of the recordings and guess what words they would encounter before listening.

Field (2008), in his book, states that there are couple of benefits of replaying the listening text during intensive listening. He also emphasizes that "...rehearing of the recorded material assists all members of the class to extend their understanding of it" (p.15). Moreover, Vandergrift and Goh (2012) mention that students can easily get motivated and understand more of the content of the spoken text when they re-listen it. Therefore, each part in the listening tests was played twice.

In the language institute where the study was conducted, every lesson takes 50-minutes and every break is 10 minutes. The first implementation, which involved the collection of the EFL students' demographic information, their EFL listening style, and their performance scores in the Listening test-1, took one class-hour. Therefore, after working in 50-minute class-hour and in 10-minute brake (i.e., totally 60 minutes), 20-minute brake was given to the students. In that time, tea and coffee were served them. After having a rest for 20 minute, the next implementation for the listening test-2 started by distributing of the last booklet. The administration of the second listening test was similar to the administration of the first listening test and the second phase of the test took 41 minutes.

As soon as completing the data collection of all three sets of data, the students in the randomly selected classrooms were asked to whether there was a volunteer to participate in a group interview. Thus, a semi-structured group interview was carried out with four voluntary participants, who had just completed to take the listening tests, from each randomly selected classroom. The interviews were held in Turkish in the researcher's office and all sessions were recorded. In the interview, participants were asked questions (see Appendix U) about the administration of the listening tests, the visuals in the tests, the presence and the absence of the visuals, the participants' preferences in terms of the visuals, the quality of the audios and videos, the topic of the recordings, and the difficulty of the test items. As can be seen from the list of the interview questions in Appendix U, the EFL students were asked different questions which were not the main focus of the study. In order to avoid leading the participants to guess what the aim of the study was and causing them to change their thoughts about the visuals and different input modes presented in the listening tests, the

researcher asked a few unrelated questions in the interview sessions. The figure 3.15 shows data collection procedure.

Regarding the scoring method of the items in the listening tests, a value of 1 was assigned for each correct answer and a value of 0 was assigned for each incorrect answer for the answers of the listening tests. Thus, as soon as completing the data collection process, all listening tests were checked by two instructors and the scores for each multiple-choice item for both listening tests were entered to Microsoft Excel (2016) program.

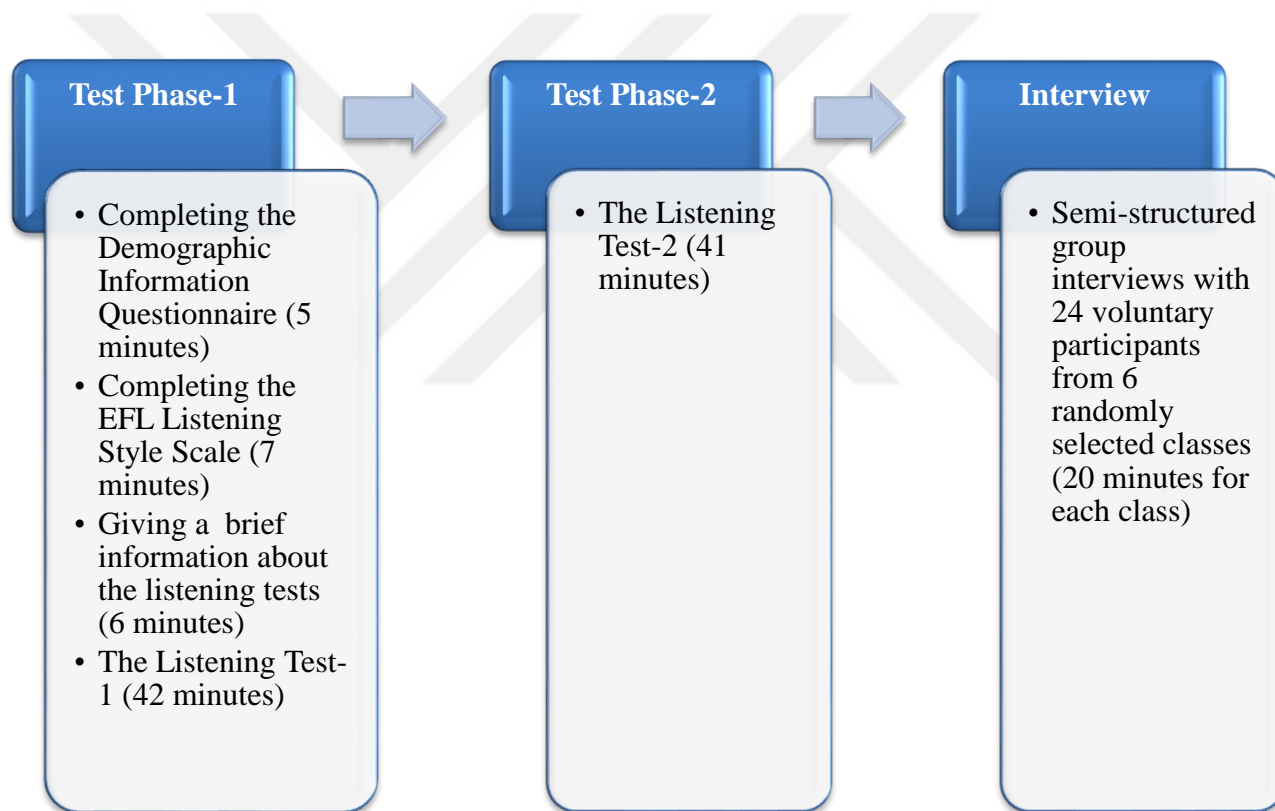


Figure 3.15. Data collection procedure

3.7. Data Analysis

All statistical analyses of the Demographic Information Questionnaire, the EFL Listening Style Scale, the Listening Test-1 and the Listening Test-2 were performed by SPSS 20.0. In the process of the item analysis including item facility and item

discrimination, and the distractor analysis, Microsoft Excel (2016) was used. Table 3.16 contains detail information about the statistical analyses and the data collection instruments used for answering the research questions of the present study.

Frequencies, means and standard deviations were used for descriptive statistics. In order to measure the reliability scores of the EFL Listening Style Scale and the Listening Test-1 and the Listening Test-2, reliability analysis was carried out. Moreover, for the validity study of the EFL Listening Style Scale, factor analysis was used.

For answering the first research question which is “Is there any significant difference among different types of input modes, namely audio-only, audio with visuals and video, in the listening tests in terms of their impact on EFL students’ academic listening performance?”, one way repeated measures-ANOVA was conducted.

Regarding the second research question, “: Is there any significant difference between different types of visuals, context and content, in the listening tests in terms of their impact on EFL students’ academic listening performance?”, a paired-samples t- test and one way repeated measures-ANOVA were performed.

Concerning the third research question, “Is there any significant effect of individual differences (proficiency level, gender, and listening style) on EFL students’ performance scores in academic listening subtests, namely audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals?”, independent sample t-test and one-way ANOVA were carried out.

Furthermore, content analysis (open-coding by Corbin & Strauss, 2008) technique was used to analyze qualitative data, which was aimed to answer the fourth research

question, obtained from semi-structured group interviews. Qualitative interview data were translated into English and then transcribed for doing the analyses. Analysis was done manually by two researchers in order to provide inter-rater reliability. The coding of the transcribed qualitative data started after analyzing some of the documents together to negotiate on the criteria, codes and themes. After completing the analysis of the qualitative data separately, it was found that out of 654 sentences, 523 sentences were agreed on and the agreement was 80%. The codes and themes, which were not agreed on, were also negotiated and an agreement was reached.

Table 3.16 Data Analyses Used for Answering the Research Questions

Research Question	Instrument	Analysis
RQ 1: Is there any significant difference among different types of input modes, namely audio-only, audio with visuals, and video, in the listening tests in terms of their impact on EFL students' academic listening performance?	Listening Tests	One Way Repeated Measures ANOVA
RQ 2: Is there any significant difference between different types of visuals, context and content, in the listening tests in terms of their impact on EFL students' academic listening performance?	Listening Tests	Paired Samples t-test and One Way Repeated Measures ANOVA
RQ 2.1: Is there any significant difference between audio with context visuals and audio with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?	Listening Tests	Paired Samples t-test

RQ 2.2: Is there any significant difference between video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?	Listening Tests	Paired Samples t-test
RQ 2.3: Is there any significant difference among audio-only, audio with context visuals, audio with content visuals, video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?	Listening Tests	One Way Repeated Measures ANOVA
RQ 3: Is there any significant effect of individual differences (proficiency level, gender, and listening style) on EFL students' performance scores in academic listening subtests, namely audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals?	Demographic Information Questionnaire EFL Listening Style Scale Listening Tests	Independent Sample t-test One Way ANOVA
RQ 3.1 Is there any significant effect of proficiency level on EFL students' performance scores in academic listening subtests?	Demographic Information Questionnaire Listening Tests	One Way ANOVA
RQ 3.2: Is there any significant effect of gender on EFL students' performance scores in academic listening subtests?	Demographic Information Questionnaire Listening Tests	Independent Sample t-test

<p>RQ 3.3: Is there any significant effect of listening style on EFL students' performance scores in academic listening subtests?</p>	<p>EFL Listening Style Scale Listening Tests</p>	<p>One Way ANOVA</p>
<p>RQ 4: What are EFL students' experiences about the presence and absence of context and content visuals provided by multi-modal input in the listening tests?</p>	<p>Semi-structured Group Interview</p>	<p>Open-coding Content Analysis</p>



CHAPTER 4

RESULTS

This chapter presents the answers of the research questions using the results of multiple analyses. In the study, four kinds of data from the demographic information questionnaire, the EFL Listening Style Scale, the Listening Test-1, the Listening Test-2 and a semi-structured group interview were analyzed in order to answer eight research questions. The following sections in this chapter provide the results of each research question.

RQ 1: Is there any significant difference among different types of input modes, namely audio-only, audio with visuals and video, in the listening tests in terms of their impact on EFL students' academic listening performance?

In order to measure whether there is a significant difference among three different input modes, audio-only, audio with visuals (involved both context and content visuals together) and videos (included both context and content visuals together), in the listening tests, the researcher compared the EFL students' audio-only scores, audio with visuals scores, and video scores using a one way repeated measures-ANOVA. The analysis was performed on the basis of the performance scores obtained from the Listening Test-1 and the Listening Test-2. The EFL participants' test performance scores from three input modes (audio-only, audio with visuals and video) in the listening tests were submitted as dependent variable with the types of input mode as independent variables. Since the number of the questions in the audio-only mode (19) was not equal to the number of the questions in the audio with visuals mode (38) and the video mode (38), the participants' test scores from each mode were converted to 100 for the comparison in the analysis. The descriptive

statistics table of three input modes from the listening tests is provided in table 4.1., and one-way ANOVA for repeated measures test results are presented in table 4.2.

Table 4.1 Descriptive Statistics of Three Input Modes from the Listening Tests

Measures	M	SD	N
Audio-Only	56.65	15.09	127
Audio with Visuals	63.67	17.46	127
Video	62.41	21.43	127

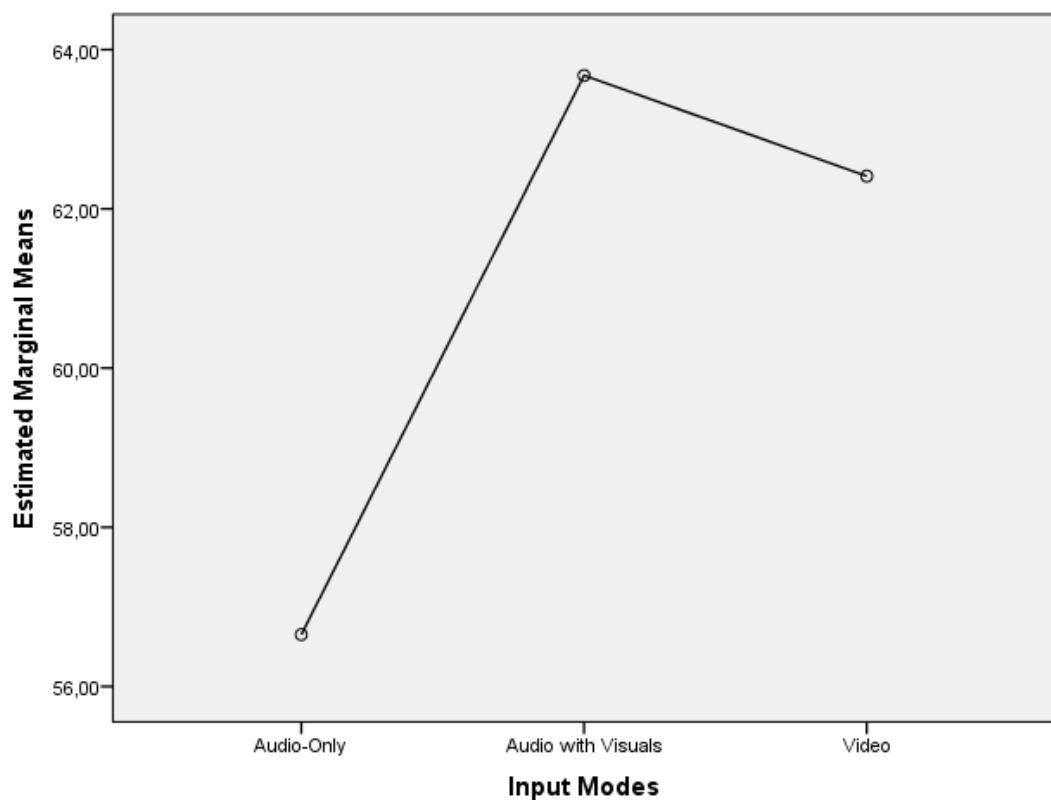


Figure 4.1 Estimated marginal means of the input modes in the listening test-1 and the listening test-2

Table 4.2 Repeated Measure ANOVA Test Result for Three Input Modes

Sources	Sum of Squares	df	Mean Square	F	p	Partial eta-square
Between Subjects	101729.7	126	807.379			
Measures	3561.23	1.626	2190.589	19.274	.000	.133
Error	23280.88	204.838	113.655			
Total	128571.81	331.626				

The Mauchly's test was used to test the sphericity assumption of the analysis. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(2) = 32.710, p = <.05$). Therefore, Greenhouse-Geisser correction for the violation of sphericity was applied to ensure assumptions. The result of the repeated measures ANOVA analysis shows that, there was a significant difference between the EFL students' three different input mode scores- namely audio-only, audio with visuals, and video- on the listening tests, $F(1.62, 204.83) = 19.27, p < .0005, \eta^2 = .133$. In order to determine which measurements were different from each other, Bonferroni multiple comparison test was conducted. As can be seen in table 4.3, post hoc analysis with a Bonferroni adjustment revealed that there was a significant difference between the EFL students', audio-only mode and audio with visuals mode. According to the analysis, there was also a significant difference between audio-only mode and video mode but there was no significant difference between audio with visuals mode and video mode.

Table 4.3 Bonferroni Test Results for Multiple Comparisons

(I) Measure 1	(J) Measure 2	Mean Difference (I-J)	SE	Sig.
Audio-Only	Audio with Visuals	-7.024*	1.261	.000
	Video	-5.760*	1.407	.000
Audio with Visuals	Audio-Only	7.024*	1.261	.000
	Video	1.264	.891	.476
Video	Audio-Only	5.760*	1.407	.000
	Audio with Visuals	-1.264	.891	.476

*The mean difference is significant at the .05 level.

RQ 2.1: Is there any significant difference between audio with context visuals and audio with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

In order to investigate the difference between the EFL students' academic listening performance in audio with context visuals and in audio with content visuals subtests, their test performance scores obtained from audio with context visuals and audio with content visuals on the listening tests were compared using a paired-samples t-test. The analysis was performed on the basis of the EFL participants' performance scores obtained from the Listening Test-1 and the Listening Test-2. The result of the analysis can be seen in table 4.4.

Table 4.4 Paired-Samples t-Test Results for Audio with Context and Content Visuals

	N	M	SD	df	t	p
				126	-1.71	.089
Audio with Context Visuals	127	11.87	3.61			
Audio with Content Visuals	127	12.32	3.64			

The result of the paired-samples t test indicated that while the EFL students' mean score for the audio with content visuals (M=12.32, SD= 3.64) was slightly

higher than their mean score for the audio with context visuals ($M=11.87$, $SD= 3.61$), there was no statistically significant difference, $t(126) = -1.71$, $p= .089$, between the EFL students' audio with context visuals scores and audio with content visuals scores on the listening tests.

RQ 2.2: Is there any significant difference between video with context visuals and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

In order to investigate the difference between the EFL students' academic listening performance in video with context visuals and in video with content visuals subtests, the researcher compared the students' test performance scores obtained from video with context visuals and video with content visuals on the listening tests using a paired-samples t test. The analysis was performed on the basis of the EFL participants' performance scores obtained from the Listening Test-1 and the Listening Test-2. The result of the analysis is presented in table 4.5.

Table 4.5 Paired-Samples t-Test Results for Video with Context and Content Visuals

	N	M	SD	df	t	p
				126	4.36	.000
Video with Context Visuals	127	12.40	4.16			
Video with Content Visuals	127	11.31	4.44			

The result of the paired-samples t test provided that the EFL students' mean score for the video with context visuals ($M=12.40$, $SD= 4.16$) was higher than their mean score for the video with content visuals ($M=11.31$, $SD= 4.44$). Therefore, there was a statistically significant difference, $t(126) = 4.36$, $p= .000$, between the EFL students' video with context visuals scores and video with content visuals scores on the listening tests.

RQ 2.3: Is there any significant difference among audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals in the listening tests in terms of their impact on EFL students' academic listening performance?

In order to measure the difference among five different input modes in the listening tests, audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals, the researcher compared the EFL students' audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals performance scores using a one way repeated measures-ANOVA. The analysis was performed on the basis of the test performance scores obtained from the Listening Test-1 and the Listening Test-2. The EFL participants' test performance scores from five input modes (audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals) in the listening tests were submitted as dependent variables with the types of input modes as independent variables. The descriptive statistics table of five input modes obtained from the listening tests is provided in table 4.6. One-way ANOVA for repeated measures test results are shown in table 4.7.

Table 4.6 Descriptive Statistics of Five Input Modes in the Listening Tests

Measures	M	SD	N
Audio-Only	10.76	2.86	127
Audio with Context Visuals	11.87	3.61	127
Audio with Content Visuals	12.32	3.64	127
Video with Context Visuals	12.40	4.16	127
Video with Content Visuals	11.31	4.44	127

Table 4.7 Repeated Measure ANOVA Test Result for Five Input Modes

Sources	Sum of Squares	df	Mean Square	F	p	Partial eta-square
Between Subjects	6796.75	126	53.94			
Measures	244.97	3.75	65.23	13.805	.000	.099
Error	2235.83	473.14	4.72			
Total	9277.55	602.89				

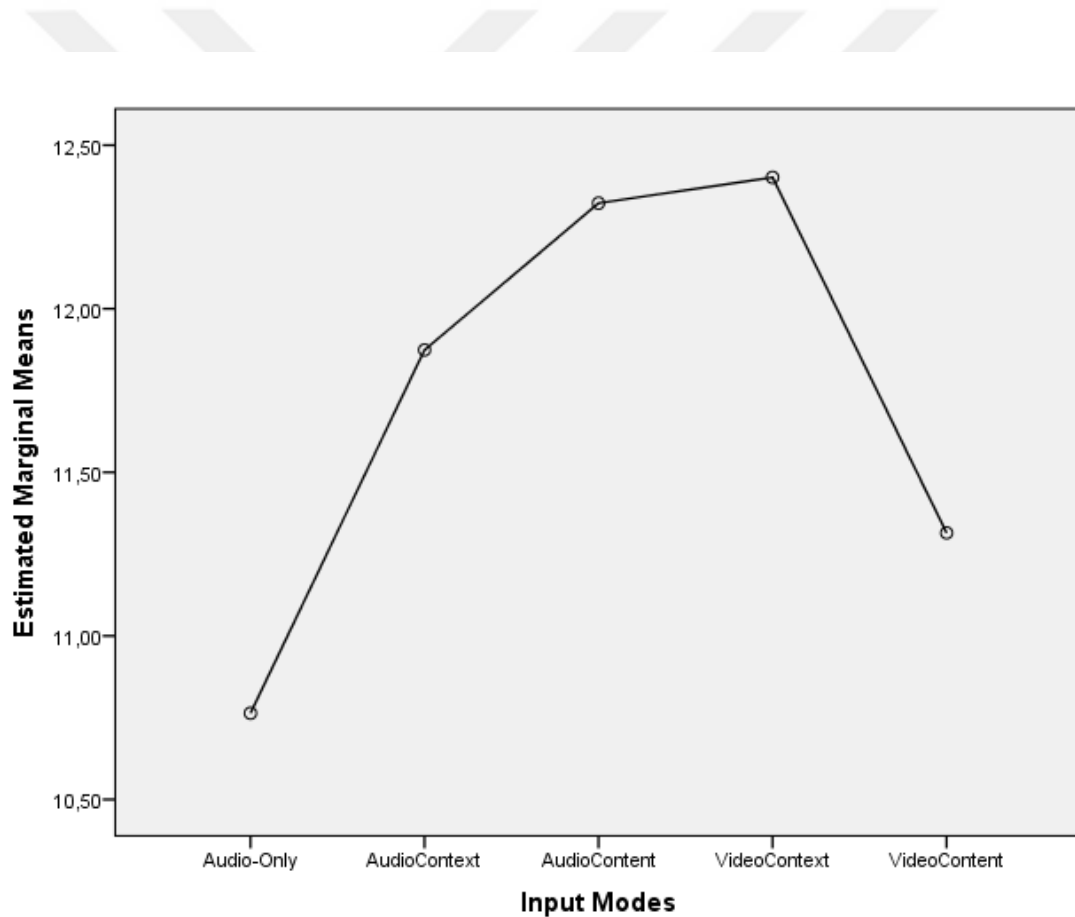


Figure 4.2. Estimated marginal means of five input modes in the listening test-1 and the listening test-2

Data is expressed as mean and standard deviation. As can be seen in table 4.6 and figure 4.2, the EFL students received the highest scores in their video with

context visuals subtest ($M=12.40$, $SD= 4.16$) on the other hand these students got the lowest test score in their audio-only subtest ($M=10.76$, $SD= 2.86$). The Mauchly's test was used to test the sphericity assumption of the analysis. The assumption of sphericity was violated, as assessed by Mauchly's test of sphericity, $\chi^2(9) = 15.17$, $p = <.05$). Therefore, Greenhouse-Geisser correction for the violation of sphericity was applied to ensure assumptions. The result of the repeated measures ANOVA analysis shows that, there was a significant difference between the EFL students' five different subtests scores- namely audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals- on the listening tests, $F(3.75, 473.14) = 13.81$, $p < .0005$, $\eta^2 = .099$. In order to determine which measurements were different from each other, Bonferroni multiple comparison test was conducted. As can be seen in table 4.8, post hoc analysis with a Bonferroni adjustment revealed that there was a significant difference between the EFL students' audio-only and audio with context visuals scores. According to the analysis, there was also a significant difference between the EFL students' audio-only and audio with content visuals scores. The analysis also reveals that there was a significant difference between the EFL students' audio-only and video with context visuals scores. Moreover, there was a significant difference between the EFL students' audio with context visuals and video with content visuals scores. In addition, another significant difference occurred between the EFL students' video with context visuals and video with content visuals scores. Finally, the last significant difference was observed between video with content visuals and audio with content visual.

Table 4.8 Bonferroni Test Results for Multiple Comparisons

(I) Measure 1	(J) Measure 2	Mean Difference (I-J)	SE	Sig.
Audio-Only	Audio with Context Visuals	-1.110*	.268	.001
	Audio with Content Visuals	-1.559*	.278	.000
	Video with Context Visuals	-1.638*	.280	.000
	Video with Content Visuals	-.551	.309	.768
Audio with Context Visuals	Audio-Only	1.110*	.268	.001
	Audio with Content Visuals	-.449	.262	.887
	Video with Context Visuals	-.528	.236	.272
	Video with Content Visuals	.559	.244	.235
Audio with Content Visuals	Audio-Only	1.559*	.278	.000
	Audio with Context Visuals	.449	.262	.887
	Video with Context Visuals	-.079	.246	1.000
	Video with Content Visuals	1.008*	.263	.002
Video with Context Visuals	Audio-Only	1.638*	.280	.000
	Audio with Context Visuals	.528	.236	.272
	Audio with Content Visuals	.079	.246	1.000
	Video with Content Visuals	1.087*	.249	.000
Video with Content Visuals	Audio-Only	.551	.309	.768
	Audio with Context Visuals	-.559	.244	.235
	Audio with Content Visuals	-1.008*	.263	.002
	Video with Context Visuals	-1.087*	.249	.000

*The mean difference is significant at the .05 level.

RQ 3.1 Is there any significant effect of proficiency level on EFL students' performance scores in academic listening subtests?

A one-way ANOVA was conducted to determine if the EFL students' academic listening performance scores on five subtests (audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals) were different for the groups with different proficiency levels.

Participants were classified into three groups based on their proficiency levels: A2

($n = 42$), B1 ($n = 42$), and B2 ($n = 43$). Data are expressed as mean and standard deviation. As can be seen in table 4.9, the EFL students in both A2 level and B1 level received the highest test scores in their audio with content visuals subtest (for A2 level, $N=42$, $M=10.11$, $SD= 3.24$; for B1 level, $N=42$, $M=11.31$, $SD= 2.82$) on the other hand these students in both A2 and B1 levels got the lowest test scores in their video with content visuals subtest (for A2 level, $N=42$, $M=7.92$, $SD= 3.74$; for B1 level, $N=42$, $M= 10.31$, $SD= 3.16$). Regarding the B2 level, while the EFL students got the highest test scores in their video with context visuals subtest ($N=43$, $M=16.51$, $SD 1.75$) and they received the lowest test score in their audio-only subtest ($N=43$, $M=12.81$, $SD= 2.01$).

As can be also seen in table 4.9, when the mean scores of each proficiency level were examined based on each subtest, the analysis revealed that B2 level students (Audio-only $M= 12.81$, Audio with Context Visual $M= 15.09$, Audio with Content Visual $M= 15.47$, Video with Context Visual $M= 16.51$, Video with Content Visual $M= 15.61$) much more increased their test scores in visually supported input modes compared to A2 level students (Audio-only $M= 8.83$, Audio with Context Visual $M= 9.78$, Audio with Content Visual $M= 10.12$, Video with Context Visual $M= 9.31$, Video with Content Visual $M= 7.93$) and B1 level students (Audio-only $M= 10.59$, Audio with Context Visual $M= 10.66$, Audio with Content Visual $M= 11.31$, Video with Context Visual $M= 11.29$, Video with Content Visual $M= 10.31$).

Table 4.9 Descriptive Statistics of Three Proficiency Levels According to the Subtests

Dependent Variable	Level	N	M	SD
Audio Only	A2	42	8.83	2.45
	B1	42	10.59	2.61
	B2	43	12.81	2.01
Audio with Context Visuals	A2	42	9.78	3.02
	B1	42	10.66	2.91
	B2	43	15.09	2.36
Audio with Content Visuals	A2	42	10.12	3.24
	B1	42	11.31	2.81
	B2	43	15.47	2.42
Video with Context Visuals	A2	42	9.31	3.57
	B1	42	11.29	2.93
	B2	43	16.51	1.75
Video with Content Visuals	A2	42	7.93	3.74
	B1	42	10.31	3.15
	B2	43	15.61	2.11

The result of the one-way ANOVA analysis revealed that there was a statistically significant difference for proficiency levels on the EFL students' academic listening performance scores. For audio-only subtest, $F(2, 124) = 30.043, p < .0005, \eta^2 = .326$; for audio with context visuals subtest, $F(2, 124) = 44.651, p < .0005, \eta^2 = .419$; for audio with content visuals subtest, $F(2, 124) = 41.558, p < .0005, \eta^2 = .401$; for video with context visuals subtest, $F(2, 124) = 72.717, p < .0005, \eta^2 = .540$; for video with content visuals subtest, $F(2, 124) = 69.581, p < .0005, \eta^2 = .529$. One-way ANOVA test results for the subtests according to proficiency levels are presented in Table 4.10.

Table 4.10 A one-way ANOVA Result Table for the Subtests According to Proficiency Level

Subtests		Sum of Squares	df	Mean Square	F	Sig.
Audio-Only	Between Groups	338.449	2	169.225	30.043	.000
	Within Groups	698.464	124	5.563		
	Total	1036.913	126			
Audio with Context Visuals	Between Groups	689.952	2	344.976	44.651	.000
	Within Groups	958.033	124	7.726		
	Total	1647.984	126			
Audio with Content Visuals	Between Groups	671.685	2	335.843	41.558	.000
	Within Groups	1002.079	124	8.081		
	Total	1673.764	126			
Video with Context Visuals	Between Groups	1180.228	2	590.114	72.717	.000
	Within Groups	1006.292	124	8.115		
	Total	2186.520	126			
Video with Content Visuals	Between Groups	1315.361	2	657.680	69.581	.000
	Within Groups	1172.041	124	9.452		
	Total	2487.402	126			

Since a significant main effect was found, Tukey post hoc analysis was designed to investigate group differences. As can be shown in Table 4.11, Tukey post-hoc tests revealed that although there was a significant difference between A2 level and B1 level, B1 level and B2 level, and A2 level and B2 level on audio-only, video with context visuals, and video with content visuals scores, there was no statistically significant difference between A2 and B1 levels on audio with context visuals ($p = .317$) and audio with content visuals subtests scores ($p = .138$). This means that EFL students' proficiency levels did not affect their performance on audio with context visuals and on audio with content visuals subtests and these two groups (A2 level and B1 level) performed similarly on audio with context visuals and on audio with content visuals subtests regardless of their proficiency levels.

Table 4.11 Tukey Post-hoc Test Results for Multiple Comparisons According to Proficiency Level

Dependent Variable	(I)	(J)	Mean	SE	Sig.
	Level	Level	Difference (I-J)		
Audio Only	A2	B1	-1.761	.518	.003
		B2	-3.980	.515	.000
	B1	A2	1.761	.518	.003
		B2	-2.218	.515	.000
	B2	A2	3.980	.515	.000
		B1	2.218	.515	.000
Audio with Context Visuals	A2	B1	-.881	.607	.317
		B2	-5.307	.603	.000
	B1	A2	.881	.607	.317
		B2	-4.426	.603	.000
	B2	A2	5.307	.603	.000
		B1	4.426	.603	.000
Audio with Content Visuals	A2	B1	-1.191	.620	.138
		B2	-5.346	.617	.000
	B1	A2	1.191	.620	.138
		B2	-4.156	.617	.000
	B2	A2	5.346	.617	.000
		B1	4.157	.617	.000
Video with Context Visuals	A2	B1	-1.976	.617	.005
		B2	-7.202	.617	.000
	B1	A2	1.976	.622	.005
		B2	-5.225	.618	.000
	B2	A2	7.202	.618	.000
		B1	5.225	.618	.000
Video with Content Visuals	A2	B1	-2.381	.671	.002
		B2	-7.676	.667	.000
	B1	A2	2.381	.671	.002
		B2	-5.295	.667	.000
	B2	A2	7.676	.667	.000
		B1	5.295	.667	.000

*The mean difference is significant at the .05 level.

RQ 3.2: Is there any significant effect of gender on EFL students' performance scores in academic listening subtests?

An independent-samples t-test was run to determine whether the EFL students' performance scores on five academic listening subtests (audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals) changed according to gender. As shown in table 4.12, the independent-samples t-test revealed that a statistically significant difference was not found on the participants' five academic listening subtests scores in terms of gender. This result shows that the participants in the study performed similarly on all five subtests regardless of their gender.

Table 4.12 Independent Samples T-Test Results for the Subtests According to Gender

Subtest	Gender	N	Mean	SD	df	t	p
Audio-Only	Male	59	10.59	3.15	125	-.623	.535
	Female	68	10.91	2.61			
Audio with Context Visuals	Male	59	11.58	3.52	125	-.863	.390
	Female	68	12.13	3.70			
Audio with Content Visuals	Male	59	11.83	4.01	125	-1.424	.157
	Female	68	12.75	3.25			
Video with Context Visuals	Male	59	11.83	4.21	125	-1.445	.151
	Female	68	12.89	4.08			
Video with Content Visuals	Male	59	11.32	4.43	125	.017	.987
	Female	68	11.31	4.47			

RQ 3.3: Is there any significant effect of listening style on EFL students' performance scores in academic listening subtests?

A one-way ANOVA was conducted to determine if the EFL students' academic listening performance scores on five subtests (audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals) was different for groups with different listening styles. The participants were classified into four groups based on their listening styles: Visual (n = 34), Spatial (n = 19), Auditory (n= 66), and Bottom-up (n = 8). Although four types of listening styles (Visual, Spatial, Auditory, and Bottom-up) were assigned on the basis of factor analysis of the EFL Listening Style Scale, 13 participants out of 127 received equal mean scores on two factors on data collection process. Since 13 participants are approximately 10 percent of all participants of the study, which is a high number, rejecting them in the analysis may cause a problem. Therefore, Professor Arif Altun, who is in the field of English Language Teaching, suggested that these 13 participants should be distributed to their one of two styles randomly. After identifying the listening styles of these 13 participants, ANOVA analysis was performed. The descriptive statistics table for the scores obtained from five different subtests according to the EFL students' listening styles is provided in table 4.13. The result of the one-way ANOVA analysis (as can be seen in table 4.14) revealed that there was no statistically significant difference for the listening styles on EFL students' academic listening performances. For audio-only subtest, $F(3, 123) = 1.067, p = .366, \eta^2 = .025$; for audio with context visuals subtest, $F(3, 123) = 1.298, p = .278, \eta^2 = .031$; for audio with content visuals subtest, $F(3, 123) = .238, p = .870, \eta^2 = .006$; for video with context visuals subtest, $F(3, 123) = .928, p = .429, \eta^2 = .022$; for video with content visuals subtest, $F(3, 123) =$

.844, $p = .472$, $\eta^2 = .020$. This means that the EFL students' listening styles did not affect their academic listening performances. This result indicates that the participants in the study were not influenced by their listening styles and performed similarly on all five subtests regardless of their listening styles.

Table 4.13 Descriptive Statistics of Listening Styles According to the Subtests

Dependent Variable	Listening Style	N	M	SD
Audio Only	Visual	34	10.47	3.06
	Spatial	19	11.79	3.11
	Auditory	66	10.57	2.74
	Bottom-up	8	11.13	2.29
Audio with Context Visuals	Visual	34	10.91	4.46
	Spatial	19	12.00	3.29
	Auditory	66	12.18	3.34
	Bottom-up	8	13.12	1.64
Audio with Content Visuals	Visual	34	12.03	3.84
	Spatial	19	12.79	3.18
	Auditory	66	12.27	3.80
	Bottom-up	8	12.88	2.69
Video with Context Visuals	Visual	34	11.71	4.44
	Spatial	19	12.90	3.79
	Auditory	66	12.39	4.21
	Bottom-up	8	14.25	3.19
Video with Content Visuals	Visual	34	10.29	4.78
	Spatial	19	11.58	4.62
	Auditory	66	11.67	4.42
	Bottom-up	8	12.13	1.88

Table 4.14 A one-way ANOVA Result Table for the Subtests According to Listening Styles

Subtests		Sum of Squares	df	Mean Square	F	Sig.
Audio-Only	Between Groups	26.289	3	8.763	1.067	.366
	Within Groups	1010.625	123	8.216		
	Total	1036.913	126			
Audio with Context Visuals	Between Groups	50.556	3	16.852	1.298	.278
	Within Groups	1597.428	123	12.987		
	Total	1647.984	126			
Audio with Content Visuals	Between Groups	9.669	3	3.223	.238	.870
	Within Groups	1664.094	123	13.529		
	Total	1673.764	126			
Video with Context Visuals	Between Groups	48.414	3	16.138	.928	.429
	Within Groups	2138.106	123	17.383		
	Total	2186.520	126			
Video with Content Visuals	Between Groups	50.170	3	16.723	.844	.472
	Within Groups	2437.232	123	19.815		
	Total	2487.402	126			

RQ 4: What are EFL students’ experiences about the presence and absence of context and content visuals provided by multi-modal input in the listening tests?

In order to support the researcher’s understanding of the EFL students’ opinions about the visuals and the input modes in the listening tests, several open-ended questions were asked to 24 participants from six different classes and three proficiency levels in a semi-structured group interview just after the administration of the listening tests. After translating the data into English and transcribing them, the participants’ qualitative data based on the questions in the interview was analyzed

through open-coding and categorical content analysis manually. In order to understand the participants' opinion with regards to the visuals and their applicability in the listening test-1 and the listening test-2, the EFL students were asked three questions (see Appendix U, Question-2, Question-4, and Question-5). The results from the analysis of these questions are given in table 4.15.

Table 4.15 Participants' Positive Opinions about the Visuals in the Listening Tests

Theme	Code	Sub-codes	f	%
Positive Opinion	Perceptions of the visuals in the tests (Total f=42)	Helpful Beneficial	18	43
		Facilitating	11	26
		Informative	5	12
		Supportive	8	19

As can be seen in table 4.15, most of the opinions about the visuals used in the listening tests were found “helpful and beneficial” (f=18) and “facilitating” (f=11) after the analysis. The EFL participants also mentioned that the visuals in the listening tests gave information (“informative” f=5) and support their understanding while answering the questions in the listening tests (“supportive”, f=8). Following excerpts illustrate these findings:

“The visuals helped me a lot especially when I didn’t know the meaning of a specific word.”

“I found the visuals very beneficial. Although I didn’t look at them in the first listening because I focused on the questions first, they helped me a lot when I listened again.”

“I took many advantages of the visuals in the tests. I can say that they gave me a lot of information about the topics.”

“The visuals were very helpful especially when I missed something during the listening. They gave me clues.”

“When I saw the visuals on the screen, I thought that I could understand more.”

“I could get a lot of information from the visuals. For example, If I didn’t understand the meaning of a word, I might understand it when I look at the visuals which is about it.”

“At least I could get the main idea of the listening text from the visuals. They made my job easier.”

“In the first time of the listening, I just focused on the questions in order to answer them correctly and then in the second time of the listening I checked my answers through visuals.”

Table 4.16 Participants' Negative Opinions about the Visuals in the Listening Tests

Theme	Code	Sub-codes	f	%
Negative Opinion	Perceptions of the visuals in the tests (Total f=14)	Confusing	3	21
		Distracting	4	29
		Noncontributing	7	50

When the table 4.16 was examined, although most of the EFL students’ opinion were positive towards the presence of the visuals and their applicability in the listening tests, few negative feedback was revealed as a result of the analysis as well. Most of the opinions about the visuals in the listening test was found “noncontributing” (f=7). Besides, some of the participants thought that the visuals distracted their attention while answering the questions (“distracting”, f=4), and few of the test-takers found the visuals “confusing” (f=3). Following statements illustrate participants’ negative opinions about the visuals and their applicability in the listening tests.

“I think the visuals did not make any contribution to my understanding.”

“According to my opinion, it was no need to put them in the tests. I didn’t get any help from them.”

“Nothing about the visuals helped me during the test.”

“Am I going to listen, am I going to look at the visuals or answer the questions. I was confused.”

“If I looked at the screen while I was listening, the images on the screen distracted my attention.”

In order to reveal the participants’ opinions with regards to the types of visuals, context and content, and their applicability and their assistance in the listening test-1 and the listening test-2, the EFL students were asked one question (see Appendix U, Question-6). The results from the analysis of this question is given in table 4.17.

Table 4.17 The Participant’ Opinions about the Assistance of the Type of the Visuals in the Listening Tests (Total f= 24)

Theme	Codes	f	%	Level of the Participants
The Assistance of the Type of the Visuals	Context	1	4	1 student from B1
	Content	13	54	3 students from A2 6 students from B1 4 students from B2
	Both	2	8	2 students from A2
	No Help	8	34	3 students from A2 1 student from B1 4 students from B2

The findings in the Table 4.17 showed that most of the EFL participants from all proficiency levels found content visuals (f=13) more helpful than context visuals (f=1) in the listening tests. It can be concluded that the participants from all levels thought that the content visuals in the listening tests made more contribution in their understanding while answering questions compare to the context visuals. Furthermore, two participants from A2 proficiency level indicated that both content and context visuals (f=2) helped them in answering questions in the listening tests. On the other hand, many students from mostly A2 and B2 levels mentioned that they did

not receive any help from the visuals (“no visual”, f=8) in the listening tests. The following excerpts prove this finding:

“The visuals helped me a lot. For example, I forgot the meaning of the word ‘earthquake’ in the test. I couldn’t remember immediately. But then, when I saw the images in the video, I just remember it.” (the participant refers to content visuals).

“I can say that the images of the listening texts about earthquake and global warming gave me a lot of information about the topics.” (the participant refers to content visuals).

“I think the visual which involved only the images of the speaker or the environment was confusing.” (the participant needs content visuals).

“The still images that involved the photos of the speakers were useless in the audio with visuals tests.” (the participant needs content visuals).

“I didn’t look at any images in the tests because I know that If I looked at them, I was lost and gave up answering the questions. I thought they distracted me.” (the participant refers to no help).

“In the EFL Listening Style Scale, I generally put a tick to the statements that refer visual listening style. But, in the listening tests, I almost did not look at any visuals on the screen. I just focused on the question papers in order to answer them correctly. If I had focused on the images, I might have missed answering questions.” (the participant refers to no help).

“The mimics and the gestures of the speaker in ‘climate change video’ helped me in the test.” (the participant refers to context visual).

“Both type of visuals helped me in different parts of the tests. For example, the images in ‘McDonald’s’ listening or the images of a man walking on a wire gave me clues.... Besides, I got some help from the gestures and the movements of a man in the listening text talking about polar bears.” (the participant refers to both type of visuals).

In order to reveal the participants’ preference of the types of visuals, context or content, in an academic listening test, the EFL students were asked a question (see Appendix U, Question7). The results from the analysis of this question is given in table 4.18.

Table 4.18 The Participant' Preference of the Type of the Visuals in a Listening Test. (Total f= 24)

Theme	Codes	f	%	Level of the Participants
The Preference of the Type of the Visuals	Context	1	4	1 student from B1
	Content	16	67	5 students from A2 5 students from B1 6 students from B2
	No Visual	7	29	3 students from A2 2 students from B1 2 students from B2

As can be seen in Table 4.18, most of the EFL students from all three proficiency levels mentioned that they prefer taking a listening test that includes content visuals (f=16). Only one student from B1 level prefers taking a listening test that involves context visuals. On the other hand, some of the participants from all three proficiency levels do not want to see any visuals (f=7) in a listening test. The following quotations support this finding:

“Since I found the visuals which consisted of the images about the content of the text more beneficial. I prefer content visuals in future listening tests.” (the participant prefers content visuals).

“In real life, while I am talking to someone, the images of the person, her mimics and gestures may help me a lot in understanding her but in the test, I I prefer visuals which involve content of the text.” (the participant prefers content visuals).

“For me, the images which involve something about the audio or the video will be better.” (the participant prefers content visuals).

“I have no problem with the images of the speaker. I can get something from the mimics and the gestures of the speaker easily, especially in videos. They give me many information and I can understand what she wants to say.” (the participant prefers context visuals).

“I prefer a listening test which does not have any visuals and images because If I look at them, my concentration is lost and then I cannot think straight anymore.” (the participant prefers no visuals).

“Listening tests which do not have any visuals are better for me.” (the participant prefers no visuals).

“In the test, visuals are not important for me because I just focused on the question papers. Visuals do not affect me.” (the participant prefers no visuals).

In order to understand the EFL student’s preference regarding the input modes, audio-only, audio with visuals and videos, during a listening test, they were asked one question (see Appendix U, Question-9). The results from the analysis of this question is given in table 4.19.

Table 4.19 The Participant’ Preference of the Type of the Input Modes in a Listening Test. (Total f= 24)

Theme	Codes	f	%	Level of the Participants
The Preference of the Type of the Input Modes	Audio-only	7	29	2 students from A2 2 students from B1 3 students from B2
	Audio with Visuals	1	4	1 student from A2
	Video	15	63	5 students from A2 5 students from B1 5 students from B2
	Doesn’t Matter	1	4	1 student from B1

Data coming from the semi-structured group interview revealed that most of the EFL students from all proficiency levels prefer taking a listening test which is delivered via either video mode (f=15) or audio-only mode (f= 7). Contrarily, only one student from A2 proficiency level wants to take a listening test which involves audio with visual mode and one student from B1 level does not care the mode of a listening test (see table 4.19). The following statements illustrate this result:

“In audio-only mode, there was no visuals.... There was a black screen. So, I could visualize what I heard and what I read. I could concentrate what I heard easily.”

“I didn’t look at any images in the listening tests because I believe that I can understand better in audio-only mode.”

“In audio with visuals, I looked at the still images on the screen but I didn’t see them. On the other hand, in videos, I watched them and I understood them.”

“I prefer videos especially the ones involve content visuals over audios with visuals.”

“In my opinion, I am good at in videos.”

“I think videos are the best. We watch something every day. I prefer it.”

“I think audio with visuals better than videos. In the videos the images go very fast, sometimes I cannot even see them but in audio with visuals I can see the images very long time.”



CHAPTER-5

DISCUSSION AND CONCLUSION

This chapter provides a discussion of the results of the present study that investigated the impact of context and content visuals in multi-modal input on academic listening comprehension. In doing so, EFL students' academic listening performance scores on five subtests which were obtained from the Listening Test-1 and the Listening Test-2, their listening styles from the EFL Listening Style Scale, their demographic information from the Demographic Information Questionnaire, and their opinions about their experiences in the listening tests, which were collected through a semi-structured group interview, were used. Firstly, the chapter begins with a discussion of the results of the study by presenting the research questions. After the discussion of the findings in respect to each research question, the chapter provides some theoretical supports with previous studies in the literature. Then, the chapter concludes with the implications for the field of second language assessment, computer-based testing, and ELT material design. At last, it provides the limitations of the study and gives some suggestions for further research.

Research Question 1

Regarding the first research question aiming at investigating the difference among three different input modes- audio-only, audio with visuals and videos-, repeated measure ANOVA test results showed that there was a significant difference among the EFL students' performance scores from three different input modes. Specifically speaking, the audio with visuals and the video modes produced significantly beneficial effects compared to the audio-only mode, and these multi modal modes improved the EFL students' performance on the listening tests. It seems

that the test-takers made more sense of what was heard when the audio text was supported by associated visuals. Presence of the visuals in the audio with visual mode improved the EFL students' understanding and enhanced their comprehension. Therefore, the EFL students' audio with visuals mode scores outperformed their audio-only mode scores in the listening tests. It may be interpreted that when the audio text was accompanied with appropriate visuals, the effect was highly facilitative during the listening tests.

The facilitative effect of the video mode is also one of the results of the study because, although no significant difference was discovered between the test-takers' audio with visuals scores and their video scores, there was a significant difference between the participants' audio-only scores and their video scores. This result is associated with the significant effect of the visuals because visuals are the main source of comprehension in videos and they provide support for understanding the verbal message. On the other hand, the possible reason for not finding a statistically significant result between the audio with visuals mode and the video mode but a statistical significant difference between the audio-only mode and the video mode or the audio-only mode and the audio with visual mode might be due to the similarity between the audio with visuals mode and the video mode. Including the visuals in the audio texts resulted in similar effects that the video modes produced. This result was expected because in the university where the study was conducted, most of the listening activities provided by the listening skill coursebooks involved audio with visuals and videos. The EFL students get in touch more with audio with visuals and video materials than audio-only materials as it was mentioned in Progosh's (1996) study. Since video mode is so commonly integrated into teaching listening skills that many students are very comfortable with its use (Wagner, 2010 b). Therefore, the

EFL students in the present study were familiar with these types of input modes and the visuals facilitated the students' understanding when they were included in both modes in the listening tests. In addition to these, although the EFL students in both audio with visuals mode and video mode took advantage of the presence of the visuals, their performance scores in the video mode were decreased compared to their scores in the audio with visuals mode. This result might be due to the appearance duration of the images on the screen because although both modes involved visuals, their lengths of display time were different. For example, while each visual appears at least for ten seconds on the screen in the audio with visuals mode, a lot of visuals may stream in ten seconds in the video mode. This event may result in missing the message coming from the visuals in the video mode and finally in answering the comprehension questions incorrectly. However, test-takers may find more time to internalize the visuals in audio with visuals mode (i.e., thinking about what information the visual carries or what the visual means together with the aural message) and may comprehend more easily the visual message associated with the aural message.

Another reason why the EFL students' video mode scores were lower than audio with visuals mode scores might be related to participants' interaction with watching the video and answering the questions simultaneously. Since the listening subtests in the study involved while listening tasks, the test-takers had to listen/watch and answer the multiple-choice questions at the same time. Therefore, doing at least two things simultaneously and the visuals streaming on the screen very fast in the video listening tests may be tiring for the test-takers and might result in the distraction of their attention, which is supported by the interview data of the study. A similar result also occurred in Ockey's (2007) research study where some participants

complained about watching the video and writing the answers concurrently. Parallel to the finding of the study, MacWilliam (1986) stated that since visuals in videos could distract students' attention, videos may reduce their comprehension.

Although the research in the literature reported mixed results about the effect of different input modes on test-takers' listening comprehension, the findings of the present study concur with the results of numerous previous research studies (Ahanghari et. al., 2013; Baltova, 1994; Hernandez, 2004; Homayoun, 2013; Maleki and Rad, 2011; Parry and Meredith, 1984; Secules, Herron, & Tomasello, 1992; Shin, 1998; Sueyoshi and Hardison, 2005; Suvorov, 2009; Wagner, 2010b; Wagner, 2013) which investigated the difference among audio-only, audio with visuals, and video modes. On the other hand, finding a statistically significant difference between the audio-only mode and the video mode contradicts the findings of some previous studies. (Coniam, 2001; Gruba, 1993; Londe, 2009; Ockey, 2007; Suvorov, 2009 and 2013).

Research Question 2

The second research question, which comprised three parts, investigated whether there was any significant difference between the different types of visuals, context and content, in the listening tests in terms of their impact on EFL students' academic listening performance. Specifically, the first part of the second research question aimed at investigating the difference between audio with context visuals and audio with content visuals in the listening tests.

Regarding the results obtained to answer the first part of the second research question (RQ: 2.1), the paired sample t-test indicated that there was no statistically significant difference between the EFL students' audio with context visuals and audio

with content visuals scores in the study. The finding showed that both context visuals (i.e., the visuals involving the images of the speaker and/or the setting) and content visuals (i.e. the visuals involving the images of the content of the message) complemented by audio mode did not differ in terms of their impact on EFL students' performance in the listening tests. The high test scores in the audio with content visuals subtest were expected. What was not expected is that EFL student's performance scores in audio with context visuals subtest were close to their scores in audio with content visuals subtest. The facilitative effect of both types of visuals- context and content- occurred, although there was no statistically significant difference. This result may be attributed to the reality that context visuals are made up of authentic elements because they mirror real life situations. In our life, most of the received verbal messages contain context visual supplementations such as the appearance of the speaker or the environment where the speech takes place. When the test-takers were faced with context visuals in the audio with context visuals subtest, these visuals that involved real-life elements (i.e., the images of the speakers, their gestures, mimics, body movements, or posture) might have caught their attention. Thus, the listeners might have tried to retrieve more information by focusing on to both the spoken input and the context visuals to a great extent.

This finding contradicts the finding of Ginther (2002), who reported that a significant interaction between stimulus type (mini-talks) and visual condition (no visual, and with visual, context and content,) was detected. Specifically, the researcher concluded that although content visuals in mini talks produced a facilitative effect, context visuals in mini talks created little debilitating effect. Furthermore, Ginther (2002) claimed that if the context visuals were not related, or

little related, to the content of the talk, test-takers might be distracted by the presence of them.

In this study, although a significant difference was not discovered between the context and content visuals on the EFL test-takers' audio with context and audio with content visuals subtests scores, the test-takers' mean scores of audio with content visuals were slightly higher than their scores in the audio with context visuals. The reason that the images in the audio with content visuals subtests produced more facilitation might be due to the fact that content visuals provide information about the content of the verbal message. That is, the visual part of the incoming message might have been completely related to the content of the verbal part of the message. This finding may be related to the strong semantic relationship between the content visuals and the audio message (Hu and Jiang, 2011). It is possible that the type of the semantic relationship between the content visuals and the verbal message in the audio with content visuals subtest is redundant (i.e. both images and the aural words carry identical meaning) according to Shriver's (1997) suggestion.

To address the second part of the second research question (RQ:2.2) aimed at investigating the difference between video with context visuals and video with content visuals in the listening tests, a paired sample t-test was conducted. The result of the statistical analysis showed that a statistically significant difference between the EFL students' video with context visuals and video with content visuals scores was discovered in the study. The finding indicated that both type of visuals (context and content), complemented by video mode, differed and had an effect on the EFL students' performance in the academic listening tests. Contrary to the finding of the first part of the second research question which revealed that the EFL students' performance scores on the audio with content visuals were higher than their scores on

the audio with context visuals, the opposite result was discovered in the second part of the second research question. In other words, the test-takers performed better in the video with context visuals subtests than they did in the video with content visuals subtest. The higher scores in the video with context visuals might be attributed to the nature of the video mode and the characteristics of the context visuals because both of them reflect real-life listening situation. Since most of us, except the visually impaired, are faced with visual cues (most of them involve context visuals) and use them to understand the verbal messages consciously or unconsciously while we are listening in real-life, we may automatically perform the same habit in a listening test, which is trying to make the incoming message meaningful, by using contextual cues, such as gestures, mimics or body movements. Therefore, test-takers might exhibit the same behavior in a listening test.

This finding contradicts the research results reached by Suvorov, (2013) who did not find any statistical difference between context and content subtest scores of the test-takers in video input mode. Furthermore, contrary to the findings of the present study, the participants in Suvorov's (2013) study scored higher in content video than they did in context video. Suvorov proposed five possible reasons for not finding a significant difference between context and content visual videos in his study as; (a) possible similarity between the context and content visuals in his videos, (b) weak semantic congruity between the visuals and the verbal message, (c) the questions did not require any support from the visuals in the listening test, (d) the type of questions (i.e., multiple-choice) used in the listening test, (e) instead of watching the videos, doing other tasks such as taking notes during the test.

With respect to the third part of the second research question (RQ: 2.3) which aimed to investigate the difference among five input modes- namely audio –only,

audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals- a repeated measure ANOVA test was conducted. The result of the analysis revealed that a statistically significant difference was found among these five different input modes. In order to find out which input modes were different from each other, a post hoc analysis was performed.

According to the post hoc analysis, although a significant difference was not discovered between the context and the content visuals on the EFL test-takers' audio with context and audio with content visuals subtests scores, a significant difference was found between the audio-only and the audio with context visuals subtests scores, and the audio-only and the audio with content visuals subtests scores. The result of the descriptive statistics showed that the mean scores of the audio with context visuals and the audio with content visuals subtests were relatively higher than the mean score of the audio-only input mode. Thus, the mean scores of both subtests compared to the audio-only subtest scores indicated that both type of visuals-context and content-facilitated the students' understanding in the listening tests. It may also be concluded that, contrary to Ockey's (2007) finding, the audio with context visuals and the audio with content visuals subtests might assess different abilities through visuals to those not assessed by the audio-only mode.

This finding is in line with the findings of Mueller (1980) who investigated the effects of context visuals on the listening comprehension process. In the study, the researcher compared the performance of two groups (low proficient and high-proficient) by providing a listening test which included audio with context visual subtest and audio-only subtest. The researcher concluded that context visuals complemented by audio input can improve the listening comprehension recall especially for beginner level of German learners. On the other hand, this finding

contradicts the result of Suvorov's (2009) study because he did not find any significant difference between audio with context visual test scores and audio-only test scores.

The result of the post hoc analysis of the present study also revealed that although any significant difference was not found between the audio-only and the video with content visuals mode, a significant difference was discovered between the audio-only and the video with context visuals mode. The reason of this result might be attributed to the nature of the videos and the characteristics of the context visuals. Because of including more real-life elements, the video with context visuals probably provide more authentic stimuli than the audio-only, audio with images, and video with content visuals modes do (as mentioned in Ockey's (2007) study). Furthermore, Wagner (2010b) mentioned that the ability to use visual information can be seen as a part of pragmatic competence because its usage is related to the incoming aural message, non-verbal information, and context of the setting. Since human beings always utilize their pragmatic competence to comprehend incoming messages in every listening situation in their life, it is quite natural to expect them to exhibit the same behavior in a second language listening test. Therefore, it is not a surprise that the video with context visual multimodal mode had the highest mean score among other input modes in the study.

This finding concurs with the finding of Wagner (2010b). The researcher concluded that the test scores of the test-takers who were in the context video mode were higher than the scores of those who were in the audio-only mode. Besides, although the finding of the present study is parallel to the finding of Suvorov (2009), who discovered a significant difference between the video mode involving context visuals and audio only mode, the mean scores of these two modes in his study

contradict the mean scores of the same modes in the present study. Specifically speaking, although the test-takers received the lowest test scores in the video with context mode, they got the highest test scores in the audio-only mode in his study. Conversely, in the present study, the EFL students' received the highest test scores in the video with context visuals mode and the lowest in the audio-only mode. Furthermore, contrary to the finding of the present study, Coniam (2000) and Gruba (1993) reported no significant difference between audio-only mode and context video mode in their studies.

In addition to these, what was surprising is that a difference between the audio-only mode and the video with content visuals mode was not found and they had a similar effect on the EFL test-takers' performance. When the verbal message was introduced with the visuals supported by only content information in the video with content visuals mode, the presence of the visuals resulted in little debilitating effect on the EFL students' test scores compared to their scores in video with context visuals. With regard to the reason, it is possible that trying to catch very fast moving images to comprehend something from the content visuals, and to understand the verbal message related to the content visuals simultaneously might have exceeded students' cognitive capacity because of a redundancy effect. This effect occurs when information is unnecessarily presented in multiple forms, then it causes working memory load to increase which creates a detrimental effect in learning (Kalyuga, Chandler, & Sweller, 1998; Sweller, 2010). Similarly, in his study, Suvorov (2013) explained that trying to interpret the content visuals in video mode might cause semantic overload because, unlike context visuals which did not carry much meaning, content visuals demanded much more mental processing by additional sensory. Therefore, the participants might have given up watching the video by closing their

visual sensory intentionally and just listened to the incoming message via auditory sensory. Another possible reason, as Wagner (2010a) claimed in his study, why test-takers might not have found the content visuals in the video mode useful - even distracting- for them to understand the verbal input or to answer the test questions. Consequently, the EFL test-takers' initial focus was on the verbal message rather than the visual information and it appears that the content visual inputs in the video did not contribute to understanding. Besides, as Cross (2011) mentioned in his study, contents of the visuals do not necessarily facilitate learners' comprehension, they sometimes inhibit attention to, and processing of verbal message.

As another finding of the last part of the second research question, a significant difference was detected between the audio with content visual mode and the video with content visual mode. Although the type of the visuals provided in both the audio and the video mode was the same, the test-takers' mean scores in the video with content visuals mode were particularly lower than their scores in the audio with content visuals mode. One of the possible reasons might be attributed to the appearance duration of the images on the screen. Since audio with content visuals subtest comprised 9-11 still images, each of which stayed approximately 10-20 seconds on the screen, the test-takers had enough time to make a meaningful connection between the verbal and the visual input. On the other side, the test-takers were probably distracted by the appearance of very fast streaming visuals in the video mode.

To address another reason why the test-takers received lower test scores in the video with content visuals mode might be due to the characteristics of the content visuals in the video mode because, contrary to the content visuals in the audio mode, the content visuals in the video mode carry a lot of information in a very short time.

The participants had to integrate a lot of messages coming from both fast streaming visuals and aural input in a very short time in order to make meaningful sense of them and to answer the questions. Therefore, many of the test-takers probably gave up watching the video on the screen and answered the questions by focusing only on listening, or they watched the video and answered the questions on open test-booklets as soon as the message was presented in the video rather than waiting until the end of the video to answer the questions. A similar result was found by Brett (1997), Gruba (1994), and Wagner (2010a).

Research Question 3

The third research question, which consisted of three parts, aimed at investigating the effect of individual differences on EFL students' performance scores in academic listening subtests. Specifically speaking, this research question tried to investigate the effects of proficiency level, gender, and listening style of EFL students on their performance scores in academic listening subtests, namely audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals.

To address the first part of the third research question (RQ: 3.1), a one-way ANOVA test was conducted to see if the EFL students' academic listening performance scores on five subtests were different for groups with different proficiency levels. The result showed that the EFL students who were in separate proficiency levels performed differently on five separate subtests. When the results related to the proficiency levels were examined thoroughly, very striking findings came out. First of all, regarding the result of the descriptive analysis, the mean scores of the test-takers at B2 (upper-intermediate) proficiency level were higher on all five

subtests than those for B1 (intermediate) and A2 (elementary) proficiency level test-takers. This was an expected result for the study because of the extensive language experience of B2 level students. This finding demonstrates that as the EFL students' English language proficiency level increases, their listening test scores increase regardless of the input mode. These findings are in accord with the findings of Ginther, (2002), Mueller (1980), Parry and Meredith (1984), and Sueyoshi and Hardison (2005) who pointed out that students with high proficiency level performed better in all modes than lower proficient learners did. Furthermore, the high proficient learners increased their test scores in the visually supported modes compared to their audio-only test scores.

Secondly, the descriptive statistics results revealed that while both A2 (elementary) and B1 (intermediate) level students outperformed in the audio with content visuals mode, these students in both levels failed in the video with content visuals mode. This result confirms the findings of the third part of the second research question, in which a significant difference was discovered between audio with content visuals and video with content visuals modes. The finding displays that since the visual type (content) in both modes was the same, the delivery format (audio or video) might have affected the result of the students' test scores and it might have caused this discrepancy. It seems that both A2 and B1 level students benefited from the presence of the content visuals in the audio with content visuals mode. Since the students in these two levels had low-level language proficiency, they might have been looking for more help from the visuals in order to make sense of the verbal message. The test-takers' inadequacy in linguistic abilities might be compensated by the support of other aids in the visually supported audio modes. Since the display duration of the content visuals on the screen in audio modes is long, the test-takers with limited linguistic

background might have enough time to integrate the information coming from the still images with the verbal information. Thus, their test performance on the audio with content visuals mode scored higher. On the other hand, trying to grasp some information, which makes the aural message meaningful, from very fast streaming content visuals in the video mode might be a distracting experience for the students with low language proficiency level because the students' limited language level may hinder the processing of the verbal message together with the visuals with ease. Consequently, their low language ability may prevent them from utilizing the content visual cues, thus, the students in A2 and B1 level may be distracted and produced the lowest test scores in the video with content visuals mode. This finding might be also attributed to the redundancy effect. That is, elaborated information which was introduced through different sources (visual and verbal) might have caused negative effects on the learners with low proficiency level. Furthermore, Vandergrift and Goh, (2012) state that the use of visual and aural at the same time, may be disturbing and challenging for the learners because the language learners continuously move their eyes between the screen, which shows visually supported listening text, and test materials. Therefore, they suggest that there should be conformity between the visuals and the content of the spoken message especially for less-skilled language learners. The finding also supports Ur's (2009) claim that listeners' memory system is so overloaded with decoding every single item coming from the foreign spoken message that listeners cannot relax and find time to take a broader view.

Thirdly, according to the result of the descriptive analysis, while B2 level students were successful in the video with context visuals mode, they failed in the audio-only mode. The EFL students' high language proficiency level made them the best in the most authentic input mode. Among five subtests, receiving the lowest test

score in the audio-only mode and the highest test score in the video with context visuals mode reveals that the high ability students are not attached to only auditory input and they need more information coming from the visual sources in order to understand the incoming verbal message better. In addition, considering the mean differences among five subtests in each level, it can be mentioned that B2 level students gained more help from the visuals than A2 and B1 level students in the listening tests. It is also clear that compared to A2 and B1 level students, only the high proficient participants received the maximum benefit from the video and they possibly integrated the visuals with the verbal information more easily because of having more foreign language experience and extensive language skills (Sueyoshi and Hardison, 2005). This result contradicts the finding of Latifi and Mirzaee's (2014) study because they did not find any significant difference between audio and video scores of high proficient learners and researchers pointed out that the low proficient learners utilize context to understand what they missed in the audio-only mode. Furthermore, contrary to the present study, Maleki and Rad (2011) found the opposite finding that the visual aids helped the lower proficient students more than the higher proficient ones.

According to the result of the descriptive statistics, which was mentioned above, although the low proficient learners produced the lowest test scores in the video with content visuals mode, the high proficient learners got the second highest test score in the video with content visuals mode. It is possible that while one group of students might have found the content visuals in video mode to be distracting, the other group might have found the visuals helpful and facilitative in the same input mode. Suvorov (2013) attributed this discrepancy to L2 test-takers' individual differences. He also reported that while some students could multitask and handle two

different channels-audio and visual- simultaneously, some of them could not multitask and their attention was distracted between different input channels. Another possible explanation might be that the high proficiency level students are more acquainted with academic mini-talk listening texts, which are in the form of context and content video, and with real-life academic class lectures (Suvorov, 2013) than the low proficiency level students. Thus, the revealing of diverse results in different proficiency levels can be attributed to the EFL students' English language proficiency level.

Finally, since the students in all three proficiency levels increased their test scores in the audio with content visuals mode compared to the audio with context visual modes, it is likely that choosing the audio with content visuals mode will be a reasonable decision for all proficiency levels when designing a listening test. Moreover, given the effect of proficiency levels on the video with context and the video with content visuals modes, it is probably the best choice to utilize video with context visuals mode for all proficiency levels while constructing a listening test because the students in all three proficiency levels got their highest test scores in this mode compared to video with content visual mode. When the students' listening test scores on five input modes according to their proficiency levels were examined, it was found that the audio-only mode was the worst mode in which students in both B1 and B2 levels performed poorly but it was the second worst mode for A2 level students.

In order to investigate the group differences, the Tukey post hoc analysis was carried out in the study. The result of this analysis also revealed that although significant differences were discovered between the three proficiency levels in almost all subtests, a significant difference was not found between A2 level and B1 level students' test scores in both audio with context and audio with content visuals modes.

The absence of a significant difference between the test scores of the students in A2 and B1 levels should not be considered as evidence that no effect exists. While the mean difference between A2 and B1 level of students' test scores was large in the audio-only mode, the mean difference of these students' test scores in the audio with context and content visuals modes was very small. Finding a significant difference between A2 and B1 levels in the audio-only mode but not in the audio with context and content visuals mode shows that both types of visuals facilitated low language ability (A2) students' work in the listening test. Therefore, A2 level students scored higher in both visually supported audio modes compared to their scores in the audio-only mode and they were as successful as B1 level students in these modes. Consequently, it is likely that the EFL students who have low language proficiency level need more support from the visuals because of inadequacy in their language skills in order to make more sense of the verbal message.

The findings of the present study related to the effect of proficiency levels on the EFL students' performances on five subtests in the listening tests contradict the findings of few research studies. For example, Gruba (1999) stated that an increase in the language proficiency led to less reliance on visual aids. Besides, Maleki and Rad (2011) pointed out that higher proficient test-takers did not benefit from visual inputs in the listening tests which involved context images accompanied by the aural message. Moreover, Wagner (2010a) mentioned that high proficient listeners might be less dependent on the visual aids and be more confident to decide not to watch the video because of their adequate linguistic competence. Finally, Batty (2015) did not find any interaction between participants' English proficiency level and the delivery format (audio and video).

To address the second part of the third research question (RQ:3.2), an independent sample t- test was conducted to see if EFL students' academic listening performance scores on five subtests were different for groups with different gender. The result of the analysis revealed that the EFL students who in two gender groups did not perform differently in five separate input modes. The result is that gender of the test-takers did not produce any effect, which is in parallel with the finding of Türkyılmaz, (2010) who did not detect any significant difference between female and male participants in an achievement test.

In order to answer the third part of the third research question (RQ: 3.3), a one-way ANOVA test was conducted to see whether EFL students' academic listening performance scores on five subtests were different for groups with listening style. The result of the analysis showed that the EFL students performed similarly in all five input modes regardless of their listening styles. The small effect size also indicates there is no significant difference between students who have different listening styles in five input modes. Although listening styles of the test-takers did not produce any effect, an interesting finding was revealed based on the result of the descriptive analysis. According to the analysis, the students who received the lowest test scores on all subtests were the students who were identified as 'visual style', on the other hand, the students who received the highest test scores on all subtests, except in audio-only mode, were the students who were identified as 'bottom-up style' on the EFL Listening Style Scale. It seems that although the bottom-up style listeners perceived the advantages of visuals, the visual style listeners did not see any benefit in taking the tests by visual supports. However, the result must be interpreted cautiously because no significant effect of the listening style on the EFL students' academic listening comprehension was found. There might be possible explanations of not

finding a significant effect of the listening styles on the EFL students' listening comprehension in five input modes. One of them might be that the listening style is not a variable that effects students' academic listening comprehension as it is found in the present study. Another possible explanation of not finding a significant effect of the listening styles may be attributed to nonhomogeneous number of the participants among each group of the listening styles in the study (that is, while 66 participants out of 127 were identified as Auditory style listeners- more than half of the participants- only 8 students out of 127 were assigned as Bottom-up style listeners). Therefore, providing a viable interpretation on the results may not be possible.

'Foreign or second language listening Style' concept has not received any attention from EFL/ESL researchers so far and there have been no studies that have investigated the effect of FL/ L2 listeners' listening style on EFL/ESL listening comprehension in the literature. Therefore, the finding of this research question was compared to the finding of a study that investigated the effect of cognitive abilities of the learners on listening comprehension. The findings contradict the result of the study of Hernandez, (2004) who found out that students perform best when the different mode of presentation (audio or video) matches their cognitive ability (verbal or spatial). For example, students with high verbal ability scored higher than students with low verbal ability in a video presentation mode of listening comprehension test.

Research Question 4:

To address the fourth research question, which revealed the EFL students' experience about the presence and absence of context and content visuals provided by multi-modal input in the listening tests, the interview data were analyzed through open-coding and categorical content analysis. Regarding the participants' opinions

about the visuals in the listening tests, 56 comments were recorded during the interviews. While 42 out of 56 comments (75 percent of the total) included participants' positive opinions about the presence of the visuals in the listening tests, only of 14 out of 56 comments (25 percent of the total) comprised the participants' negative opinions. According to the analysis, most of the test-takers found the visuals helpful and beneficial (43%), facilitating (26%), informative (12%), and supportive (19%). These results support the findings of the first research question which revealed that the students' test scores in the audio visual mode and the video mode were higher than their scores in the audio-only mode. The findings are also parallel to the findings of similar studies (Ockey, 2007; Gruba, 2007; Suvorov, 2013), which reported that visuals accompanied by either video or audio helped most of the students to construct meaning in listening tests. On the other hand, similar to some studies (e.g., Coniam, 2001; Ockey, 2007) few test-takers in these studies reported how visuals in different input modes distracted their attention, and thus, they did not make use of them.

Regarding the assistance of the visual types- whether context or content- in different input modes in the listening tests, majority of the students (54%) declared that content visuals were more helpful than context visuals (4%) in both audio and video modes. While 8% of the participants reported that both types of visuals helped them a lot in understanding the listening texts, 34% of them stated that both types of visuals did not help in comprehending the listening texts delivered via different input modes. This result does not support the result of the third part of the second research question(RQ:2.3), which revealed that students produced higher test scores in the video with context visuals subtest. Although only 4 % of the participants in the semi-structured group interview stated that the context visuals helped them in the listening tests, most of them scored higher in the video with context visual mode. This result

also does not support the finding of the first part of the third research question (RQ:3.1). Although it demonstrated that B2 level students were the most successful in the video with context visual mode based on the listening test scores, none of the B2 level students mentioned anything about the assistance of the context visuals in the interview. The finding obtained from the interview data demonstrates that the test-takers were not aware of how much the presence of the context visuals in the listening tests helped them. On the other hand, the interview data concerning the assistance of types of visuals indicate that the test-takers are very well aware of the help of the content visuals in the tests. The result of the interview data strengthens the result of the first part of the second research question (RQ:2.1) and the first part of the third research question (RQ:3.1), which revealed that students, especially the ones in A2 and B1 level, got the highest test scores in the audio with content visuals mode. This finding is parallel to the findings of Suvorov's (2013) study, who mentioned that participants found content visuals more helpful on content video subtests than context visuals on context video subtests for answering comprehension questions.

To address the preference of types of visuals, the interview data show that majority of the participants (67%) preferred a listening test which included content visuals whereas only 4% of them preferred context visuals. This finding strengthens the result of the first part of the third research question (RQ:3.1) one more time because most of the students who mentioned the preference of content visuals were the students in A2 and B1 level. Moreover, these students received the highest scores in the audio with content visuals mode. However, it is very interesting to find that, in the interview, 29% of the participants preferred to take the listening tests without any visuals despite receiving their lowest test scores from the audio-only mode. Since the qualitative parts of the research studies in the literature included only context type of

visuals rather than content visuals (e.g., Coniam, 2001; Ockey, 2007; Wagner, 2010a; except Suvorov, 2013), all participants in these studies preferred the context visuals compared to non-visual input mode. Yet, when the content visuals were also included as an independent variable in the design, participants had to think about which types of visuals served best in the listening tests. Thus, the students chose the one that was semantically richer in the present study and in Suvorov's (2013) study.

With regard to the preference of the type of input modes, although majority of the participants (63%) were in favor of video mode, 29% percent of them preferred audio-only mode. This finding supports the result of the third part of the second research question (RQ:2.3). This research question revealed that the students produced the highest test scores on the video with context visuals mode. On the other hand, it is very unusual to find that almost one third of the participants of the interview chose to listen to a text without visuals even though they got their lowest test scores in this input mode. Moreover, it is very interesting to discover from the verbal data that although most of the students (i.e., the students in both A2 and B1 level together) received their highest test scores on the audio with content visuals mode, only one out of 24 students preferred audio with visuals mode as the preference of type of input mode in the interview. The findings of the last part of the interview data were similar to the findings in some research studies, such as Coniam, (2001); Londe, (2009); Ockey, (2007); Parry and Meredith, (1984); Progoosh, (1996); Wagner, (2010a), because the verbal data in these studies indicated that L2 students felt that they did better on video mode. Moreover, L2 students stated that the video input did not distract their attention, on the contrary, increased their motivation, and finally they provided more assistance in comprehending the listening texts. Therefore, since these

students developed more positive attitudes towards the inclusion of the video mode in the listening tests, they preferred to take a listening test on a video mode.

Conclusion

The overall purpose of the present study is to investigate the impact of context and content visuals in multi-modal inputs on English as a foreign language students' academic listening comprehension. The study specifically aimed to examine: a) whether there exists a difference in EFL students' performance on five listening subtests, namely, audio-only mode, audio with context visuals mode, audio with content visuals mode, video with context visuals mode, and video with content visuals mode, b) whether individual differences, that is English language proficiency level, gender, and EFL students' listening styles, affect students' academic listening comprehension. In doing so, EFL students' demographic information collected by a demographic information questionnaire, their listening styles identified by the EFL Listening Style Scale, their listening test scores in five different input modes (audio-only, audio with context visuals, audio with content visuals, video with context visuals, and video with content visuals) collected by two listening tests, and finally their opinions about the listening tests collected by semi-structured group interview, were analyzed. Based on the results of the research questions, the present study suggests five important conclusions as:

Firstly, the visuals (regardless of their types), accompanied by audio mode produced the most significant effect on the EFL students' academic listening comprehension compared to the audio-only mode and the video mode because the presence of the visual input improved the EFL students' understanding and enhanced their listening comprehension. Therefore, the EFL students received their highest test scores on the audio with visuals mode.

Secondly, regarding the categorization of the visuals as context and content in audio visual mode, both the context visuals and the content visuals complemented by audio mode did not differ in terms of their impact on the EFL students' performance in the academic listening tests because these students produced almost similar results on both audio with content visuals and audio with context visuals mode. Since the facilitative effect of both types of visuals-context and content- complemented by audio mode occurred in the listening tests, a statistically significant difference was not discovered between the EFL students' test scores on the audio with content visuals and the audio with context visuals mode. Thus, it can be concluded that although the EFL students benefit similarly from the content and context visuals accompanied by audio, they perform more satisfactorily on the audio with content visuals mode in the listening tests.

Thirdly, although both types of visuals-context and content- complemented by video mode had an effect on the EFL students' performance scores in the listening tests, the students' test scores on the video with context visuals mode were higher and statistically significant than their scores on the video with content visuals mode. Therefore, it can be concluded that the EFL students benefit more from the presence of the context visuals than the content visuals on the video mode.

Fourthly, as for the effect of English language proficiency level on EFL students' listening performance, the statistical analyses indicated that the students with lower level of proficiency were more successful on the audio with content visuals mode whereas they performed poorly on the video with content visuals mode in the listening tests. However, while the students with higher level of proficiency received their highest test scores on the video with context visuals mode, they did not score well on the audio-only mode in the listening tests. In addition, based on the

mean differences between the input modes in each level, the students with higher level of proficiency (B2) gained more help from the visuals than the students with lower level of proficiency (A2 and B1) in the listening tests. Furthermore, while the students in all proficiency levels increased their listening test scores both on the audio with content visuals mode and the video with context visuals mode, the students in all proficiency levels except A2 level produced their lowest test scores on the audio-only mode. Thus, it can be concluded that the audio with content visuals mode and the video with context visuals mode can serve best for students in all proficiency levels.

Finally, as for the effect of gender and listening styles on EFL students' listening performance, the statistical analysis revealed that the EFL students, who were in different gender groups and who had different listening styles performed similarly in five different input modes. Thus, it can be concluded that gender and listening styles do not produce any effect on the EFL students' academic listening comprehension.

To sum up, the findings of the present study suggest that the inclusion of nonverbal information, whether context or content, via audio and video mode in an academic listening comprehension test has a positive effect on the EFL students' academic listening comprehension. Moreover, although the EFL students' language proficiency levels have different effects on their listening performance to varying degrees, their gender and listening styles do not have any effect on their listening performance. All in all, based on Mayer's Cognitive Theory of Multimedia Learning, the findings of the present study demonstrate that the EFL listeners present more understanding from words and visuals together than words alone. Thus, the conclusions of the study will make various contributions to several fields.

Implications

The findings of the present study have valuable practical and pedagogical implications for English as a foreign and second language instructors, English language teaching material designers and publishers, and foreign and second language test designers. The implications are discussed regarding the integration of context and content visuals complemented by audio and video mode into English language listening instruction, English language teaching materials, and English language testing.

Based on the findings of the present study, both context and content visuals accompanied by audio and video mode in the listening comprehension tests have a facilitative effect on EFL students' academic listening comprehension. With regard to the content visuals, participants perform best on audio with content visuals mode in a listening test. Although the effect of context visuals is not as high as the effect of content visuals, context visuals in audio mode also help EFL listeners to understand the incoming message. Therefore, language teachers in listening skill courses should give more place to audio with content and context visuals mode rather than audio-only mode in their instruction. In terms of finding appropriate listening texts with visuals, the Internet helps teachers providing huge amounts of sources.

In addition, in terms of developing English language listening materials, textbook publishers should include listening passages with audio with context visuals in their listening skill course books as well because almost all of the listening activities in these books comprise either content audio or content video mode (e.g., you can see in the series of *Unlock*, *Skills*, and *Pathways Listening and Speaking* course books). In addition, today, the biggest dilemma between teaching and testing

the listening skill comprehension is that although the listening skill text books involve passages that include visuals in audio or video mode, testing of it is done via only audio-only mode in the classrooms. As Scholz, (1993) mentioned that test items and test specifications should be arisen from the same needs analysis and syllabus used for designing teaching materials in order to provide a sync among learning, teaching, and testing. In addition, if a coursebook includes video-mediated listening activities, test designers should include video mode in listening tests in order to meet construct and content validity. (Bachman, 1990 as cited in Progosh, 1996; Scholz, 1993).

Furthermore, when the findings are taken into consideration from the perspective of language test construction, the implementation of visuals either content or context with audio mode in testing should be inevitable. When the current high stake tests are examined, most of them do not utilize the visuals in their listening parts. However, the listening parts of the TOEFL exam only comprise audio with context visuals or audio with content visuals that accompanies context visuals together as well. When the characteristics of the TOEFL listening comprehension tests are examined, it can be seen that majority of the visuals are context related (Ginther, 2002). Under the light of the findings of the present study, it can be suggested that test designers should integrate visuals into listening comprehension tests- even into the international ones as well- and if they would like to use visuals with an audio mode in a listening test, the integration of content visuals with audio mode will be the best decision for EFL students in all proficiency levels, especially for the low level students.

Regarding the context visuals, while the high level participants performed best on the video with context visuals mode in the listening tests, the low level ones performed well on this mode as well. Today, English language teachers can bring

different listening sources, mostly authentic ones available on the Internet, in their classrooms. In addition, different universities or institutions provide downloadable video lectures, most of them in context video mode, on the Internet. In this input mode, the visibility of the speakers' gestures, mimics, and movements may help listeners construct meaning from the verbal message. Thus, L2 students should benefit from these resources, which involve real-life features, in terms of improving their listening comprehension through video with context visuals mode.

Given that context visuals accompanied by video mode can improve EFL students' listening comprehension, especially for the high ability students, ELT material designers and publishers should not stay away from this fact. When most of the listening skill textbooks were investigated, it was revealed that almost all of the video mode listening passages comprised content visuals. For example, if the topic was about festivals around the world, the video included different visuals of festivals popular around the world. However, this study discovered that the video with context visuals helped the EFL students more than the video with content visuals did. Therefore, designing listening materials involving context visuals (that is, showing the still images of the speaker or the environment where the speech takes place) for listening skill instruction and including not only content videos but also context videos in listening skill textbooks should be considered seriously by ELT material designers and publishers.

Nowadays, with regard to test design, the biggest conflict is that there is a great incompatibility between the input modes of the listening texts in the tests and the input modes of the listening texts in the textbooks. Although the books include both content audios and content videos about varying topics for each unit, their exam materials involve only audio-only mode listening passages, as it was mentioned

above. Thus, an inconsistency arises between what to teach and what to assess. For this reason, the test developers, who would like to use context or content visuals accompanied by audio or video mode in a listening test, should be very careful about choosing appropriate visuals and they must determine clearly what to assess. It can also be suggested that if the test designers would like to use context visuals in a listening test, video mode will be the best decision for EFL students in all proficiency levels, especially for the high level learners.

Limitations

Although the results of the present study provide valuable findings related to the effects of visuals in EFL listening comprehension tests, these findings should be considered and interpreted in the light of their limitations. The present study has five limitations. One of the limitations is related to the type of the test task in the listening tests because the test tasks in the study consist of multiple-choice questions. As, (Hearst, 2000) mentioned that the questions that demand comprehensive answers are more effective than multiple-choice type questions. Moreover, Flowerdew and Miller (2013) stated that multiple choice questions test listeners' both listening and reading skills and if learners are weak in reading, this situation will affect learners' test scores negatively not because of their weakness in listening but because of their inability to understand what they read. Although the effects of other skills such as speaking and writing were eliminated by choosing multiple-choice questions for the listening tests, the effect of the reading skill was inevitable especially on the test-takers with low proficiency level.

Another limitation is related to the duration of the listening tests and the number of the subtests. In order to increase the reliability of the listening test, the

number of the questions was increased by adding one more listening test, which was similar to the first one. Since the duration of one listening test was approximately 42 minutes, only the administering process of these two tests took approximately 84 minutes. Moreover, since there were ten subtests in these two listening tests, there might have been difficulties in concentrating on the input and the participants might have been bored or lost their attention during the tests. As Buck (2001) mentions in his book, longer texts in listening can be exhausting for listeners. Therefore, it could have been better to reduce the number of the tests and subtests in order to decrease the duration of the listening tests and to keep the test-takers' attention alive.

The text type in the listening tests may be another limitation of the study because all subtests provided in five different input modes included only academic mini-talks. Nowadays, L2/FL learners who want to improve their level of English may benefit from different listening sources, mostly authentic ones, available on the Internet, or on TV. Through these authentic texts, language learners can exploit different text types, such as narratives, reports, interviews, or conversations, and familiarize with different themes and topics (Vandergrift and Goh, 2012). Since the variety of text types (dialogue, discussion, or lecture) accompanied by context and content visuals in audio or video mode may produce different results, the effects of different text types should be examined in other research studies.

Another important limitation of the study is related to the EFL Listening Style Scale. Not finding a significant effect of listening styles on the students' listening test scores in either mode attracted our attention. As it was explained in discussion section above, there might be possible explanations of this result. One of them might be that the listening style is not a factor that affects students' academic listening comprehension and another explanation may be attributed to nonhomogeneous

number of the participants among each group of the listening styles (that is, while 66 students out of 127 were identified as Auditory style listeners, only 8 students out of 127 were assigned as Bottom-up style listeners) in the study. Listening styles, like learning styles, are somehow related to learners' personality. Therefore, the researcher assumes that the result of the study related to the effect of EFL students' listening style might have been different, if it had been applied to a larger number of population.

The study did not involve EFL students' viewing behavior while watching context and content visuals accompanied by audio and video mode. During the test, participants were not forced to look at the screen while the videos or the audios with visuals were being played. Based on the interview data, some students ignored to watch the visuals because they mentioned that the appearance of them on the screen distracted their attention. Therefore, we are not exactly sure about the relation between helpfulness of the visuals on test scores and the participants' viewing behavior of them. It could have been better to collect each test-taker's viewing behavior data via observation (see in Ockey, 2007; Wagner, 2007, 2010a) or eye-tracking technology (see in Suvorov, 2013) in each mode in order to see how each participant varies in their ability to use visual information and how this ability affects their test scores.

Finally, the rationale behind this research study is the fact regarding how we listen in real life situations. However, since this study was conducted in the EFL context, most of the real life verbal interaction takes place only in students' native language not in their L2. The EFL students in this study generally use English in language learning situations in their school environment. Therefore, the findings of this study are limited to generalize only EFL university level population.

Recommendations for Further Research

Investigating the impacts of context and content visual accompanied by audio and video modes and the effects of individual differences on EFL students' academic listening comprehension have answered a lot of questions, but has also raised a number of questions to be answered in the future. First of all, in order to discover more generalizable results and to confirm the findings of this study, more replication studies are needed in either EFL or ESL contexts. Especially, we need to see the results of other research studies that will use the EFL listening Style Scale, which was developed for the present study. For this reason, this scale can be also applied to other EFL contexts with a larger sample size.

Secondly, the present study used multiple-choice questions as test task type and mini talks as text type in the listening tests. The studies in the future can integrate different test tasks (e.g., short answers or writing a summary) and different text types (e.g., lecture or dialogue) into the same research design that the present study used in order to investigate whether the results would be confirmed.

Thirdly, all few studies, including the present study, did not investigate the effects of visual information on young learners. Therefore, the studies in the future can examine the role of both context and content visuals on young EFL learners' listening comprehension in order to reveal whether visual information produces the same effects for young learners of English.

Fourthly, in the present study, only the listening texts comprised visual information not the multiple-choice questions in the test booklets. The effect of the inclusion of the same visual information, which appears in the listening test, in the questions of the test can be examined by future researchers. Thus, the investigation of

the effects of visuals in a listening comprehension test can be carried to another level in the future.

Finally, because of the inefficiency of the technological tools, the listening tests in the study were shown on a large Smartboard screen in each classroom and the test-takers answered the questions by using pen and paper. The researcher wishes that these tests would be given as computer based listening comprehension tests so that every student can take these listening tests individually via computers. Future studies can integrate the same data collection instrument into computer-based technology- this can also be configured as a computer adaptive test- thus, they can compare the results with the present study. In addition, using computer technology eases the utilizing of eye-tracking technology so that researchers can also investigate the test-takers' viewing behavior easily.

REFERENCE

- Ahanghari, S., Mozaheb, M. A., & Mirzapour, F. (2013). Testing listening in an EFL context via news videos: an exploratory study. *International Journal of Science and Technology*, 3(2), 36-42.
- Alderson, C. (1984). Reading in a Foreign Language: a Reading Problem or a Language Problem? in Alderson, J.C. and Urquhart, A.H. (eds), *Reading in a Foreign Language*. London: Longman, pp. 1–27.
- Alderson, C. J., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: CUP.
- Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge, UK: Cambridge University Press.
- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: OUP.
- Baltova, I. (1994). The impact of video on the comprehension skills of core French students. *Canadian Modern Language Review*, 50(3), 507-531.
- Batty, A. O. (2015). A comparison of video-and audio-mediated listening tests with many-facet Rasch modeling and differential distractor functioning. *Language Testing*, 32(1), 3-20.
- Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL 2000 listening framework*. Princeton, NJ: Educational Testing Service.

- Bodie, G. D., & Worthington, D. L. (2010). Revisiting the listening styles profile (LSP-16): A confirmatory factor analytic approach to scale validation and reliability estimation. *The International Journal of Listening*, 24, 69-88.
- Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14, 14-31.
- Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, 25(1), 39-53.
- Brindley, G. (1998). Assessing listening abilities. *annual Review of Applied Linguistic* 18, 171-191.
- Brown, G. (1977). *Listening to Spoken English*. London: Longman.
- Brown, G., & Yule, G. (1983). *Teaching the Spoken Language*. Cambridge: Cambridge.
- Brown, J. D. (2005). *Testing in Language Programs: A Comprehensive Guide to English Language Assessment*. New York: McGraw-Hill College.
- Brown, S. (2011). *Listening Myths*. Ann Arbor: University of Michigan Press.
- Bruno, J. E., & Dirkzwager, A. (1995). Determining the optimal number of alternatives to a multiple-choice test item: An information theoretic perspective. *Educational and Psychological Measurement*, 55(6), 959-966.
- Buck, G. (1991). The testing of listening comprehension: an introspective study. *Language Testing*, 8(1), 67-91.
- Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.

- Buck, G., & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing: Examining attributes of a free response listening test. *Language Testing* 15/2, 119-57.
- Büyüköztürk, Ş. (2004). *Sosyal Bilimler İçin Veri Analizi El Kitabı: İstatistik, Araştırma Deseni, SPSS Uygulamaları ve Yorum*. Ankara: Pegem Yayınları.
- Chiang, C. S., & Dunkel, P. (1992). The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *TESOL quarterly*, 26(2), 345-374.
- Coakley, C. G., & Wolvin, A. D. (1986). Listening in the native language. In B. H. Wing, *Listening, reading, and writing: Analysis and application*. (pp. 11-42). Middlebury, VT: The Conference.
- Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System*, 29(1), 1-14.
- Conrad, L. (1985). Semantic versus syntactic cues in listening comprehension. *Studies in Second Language Acquisition*, 7(01), 59-69.
- Corbin, J. M., & Strauss, A. L. (2008). *Basics Qualitative qesearch: Techniques and procedures for developing grounded theory*. Sage Publications, Inc.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

- Cross, J. (2011). Comprehending news videotexts: The influence of the visual content. . *Language Learning & Technology*, 15(2), 44-68.
- Cumming, A. (2013). Assessing integrated writing tasks for academic purposes: Promises and perils. *Language Assessment Quarterly*, 10(1), 1-8.
- Douglas, D. (1988). Testing listening comprehension in the context of the ACTFL proficiency guidelines. *SSLA*, 10, 245-261.
- Dunkel, P., Henning, G., & Chaudron, C. (1993). The assessment of an L2 listening comprehension construct: A tentative model for test specification and development. *The Modern Language Journal*, 77(2), 180-191.
- Farhady, H., & Shakery, S. (2000). Number of options and economy of multiple-choice tests. *Roshd Foreign Language Teaching Journal*, 14, 57.
- Field, J. (2008). *Listening in the Language Classroom*. Cambridge: Cambridge University Press.
- Flesch, R. (1948). A New Readability Yardstick. *Journal of Applied Psychology*, 32(3), 221-233.
- Flowerdew, J. (1994). *Academic Listening: Research perspectives*. Cambridge: Cambridge University Press.
- Flowerdew, J. (1994). Research of relevance to second language lecture comprehension: An overview. In J. F. (Ed), *Academic listening: Research perspectives* (pp. 7-29). Cambridge, UK: Cambridge University Press.
- Flowerdew, J., & Miller, L. (1992). Student Perceptions, Problems and Strategies in Second Language Lecture Comprehension. *RELC Journal* 23 (2), 60-80.

- Flowerdew, J., & Miller, L. (2010). Listening in a second language. In A. D. Wolvin, *Listening and human communication in the 21st century* (pp. 158-177). Oxford,UK: Wiley-Blackwell.
- Flowerdew, J., & Miller, L. (2013). *Second Language Listening Theory and Practice*. New York: Cambridge Language Education.
- Fornell, C., & Larcker, D. (1981). Structural equation models with unobservable variables and measurement errors. *Journal of Marketing Research*, 18 (2) , 39–50.
- Fraenkel, J. R., & Wallen, N. E. (2006). *How to Design and Evaluate Research in Education*. New York: Mc Graw-Hill.
- Freedle, R., & Kostin, I. (1999). Does the text matter in a multiple-choice test of comprehension? the case for the construct validity of TOEFL's minitalks. *Language Testing*, 16(1), 2-32.
- Gadermann, A. M., Guhn, M., & Zumbo, B. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. . *Practical Assessment, Research & Evaluation*, 17(3),, 1-13.
- George, D., & Mallery, P. (2001). *SPSS for Windows: Step by Step*. . USA: Allyn & Bacon.
- Gilakjani, A. P., & Ahmadi, M. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research*, 2(5), 977-988.
- Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19(2), 133-167.

- Goh, C. M. (2000). ' A cognitive perspective on language learners' listening comprehension problems'. *System* 28, 55-75.
- Graham, S. (2006). Listening comprehension: The learners' perspective. *System*, 34, 165-182.
- Griffiths, R. (1992). Speech rate and listening comprehension: Further evidence of the relationship. *TESOL quarterly*, 26(2), 385-390.
- Gruba, P. (1993). A comparison study of audio and video in language testing. . *JALT journal*, 15(1), 85-88.
- Gruba, P. (1994). Design and development of a video-mediated test of communicative proficiency. *JALT Journal*, 16(1), 25-40.
- Gruba, P. (1997). The role of video media in listening assessment. *System*, 25(3), 335-345.
- Gruba, P. A. (1999). The role of digital video media in second language listening comprehension. PhD thesis, Department of Linguistics and Applied Linguistics, University of Melbourne, from <http://hdl.handle.net/11343/38784>
- Gruba, P. (2004). Understanding digitized second language videotext. *Computer Assisted Language Learning*, 17(1), 51-82.
- Gruba, P. (2007). "Decoding visual elements in digitised foreign newscasts" . 2007 *ICT: Providing choices for learners and learning. Proceedings ascilite Singapore*, (pp. 347-356).

- Haladyna, T. M., & Downing, S. (1993). How many options is enough for a multiple-choice test item? *Educational and Psychological Measurement*, 53(4), 999-1010.
- Hansen, C., & Jensen, C. (1994). Evaluating lecture comprehension. J. Flowerdew içinde, *Academic Listening: research perspectives* (s. 241-268). Cambridge: Cambridge University Press.
- Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language Culture and Curriculum*, 13(2), 137-153.
- Hearst, M. A. (2000). Hearst, M. A. (2000). The debate on automated essay grading. *IEEE Intelligent Systems and their Applications*, 15(5), 22-37.
- Hernandez, S. S. (2004). *The effects of video and captioned text and the influence of verbal and spatial abilities on second language listening comprehension in a multimedia learning environment*. New York: New York University, School of Education.
- Homayoun, M. R. (2013). The impact of visual aids on listening comprehension tests of intermediate EFL learners. *The Iranian EFL Journal*, 19, , 83-93.
- Hu, G. G., & Jiang, N. (2011). Semantic integration in listening comprehension in a second language: Evidence from cross-modeling priming. In P. Trofimovich, & K. McDonough, *Applying priming methods to L2 learning, teaching and research: Insight from psycholinguistics* (pp. 199-218). Amsterdam, Netherlands: Jon Benjamins Publishing Company.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.

- Jensen, E. D., & Vinther, T. (2003). Exact repetition as input enhancement in second language acquisition. *Language Learning*, 53(3), 373-428.
- Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of expertise and instructional design. Human factors. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(1), 1-17.
- Kelch, K. (1985). Modified input as an aid to comprehension. *Studies in second language acquisition*, 7(01), 81-90.
- Kinsella, K. (1993). Perceptual learning preferences survey. J. M. Reid içinde, *Learning styles in the ESL/EFL classroom* (s. 221-238). Boston, MA: Heinle & Heinle.
- Krashen, S. D. (1985). *The Input Hypothesis: Issues and Implication*. Harlow and New York: Longman.
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London: Longman.
- Latifi, M., & Mirzaee, A. (2014). Visual support in assessing listening comprehension: Does it help? *International Journal of Research Studies in Educational Technology*, 3(2), 13-20.
- Londe, Z. C. (2009). The effects of video media in English as a second language listening comprehension tests . *Issues in Applied Linguistics*, 17(1), 41-50.
- Long, D. (1990). What you don't know can't help you. An explanatory study of background knowledge and second language listening comprehension. *Studies in second language acquisition*, 12(01), 65-80.

- Lord, F. M. (1977). Optimal number of choices per item—A comparison of four approaches. . *Journal of Educational Measurement*, *14*(1), 33-38.
- Lynch, T. (1998). Theoretical perspectives on listening . *Annual Review of Applied Linguistics*, *18*, 3-19.
- Lynch, T. (2006). Academic listening: Marrying top and bottom. In E. Uso-Juan, & A. Martinez-Flor, *Current trends in the development and teaching of the four language skills* (pp. 91-110).
- Lynch, T. (2009). *Teaching Second Language Listening*. New York: Oxford University Press.
- Lynch, T. (2011). Academic listening in the 21st century: Reviewing a decade of research. *Journal of English for Academic Purposes*, *10*(2), 79-88.
- MacWilliam, I. (1986). Video and language comprehension. *ELT Journal*, *40*, 131-135.
- Maleki, A., & Rad, M. S. (2011). The effect of visual and textual accompaniments to verbal stimuli on the listening comprehension test performance of Iranian high and low proficient EFL learners. *Theory and Practice in Language Studies*, *1*(1), 28-36.
- Markham, P., & Latham, M. (1987). The Influence of Religion-Specific Background Knowledge on the Listening Comprehension of Adult Second-Language Students. *Language learning*, *37*(2), 157-170.
- Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? . *Educational Psychologist*, *32* (1), 1-19.

- Mayer, R. E. (2005a). Introduction to multimedia learning . In R. E. Mayer, *The Cambridge Handbook of Multimedia Learning* (pp. 1-16). New York: Cambridge University Press.
- Mayer, R. E. (2005b). Cognitive theory of multimedia learning. In R. E. Mayer, *The Cambridge Handbook of Multimedia Learning* (pp. 31-48). New York: Cambridge University Press.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Mecartty, F. H. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning, 11*, 323-348.
- Mendelsohn, D. J. (1994). *Learning to listen: A strategy-based approach for the second-language learner*. San Diego, CA: Donimie Press.
- Mendelsohn, D. J. (1998). Teaching listening. *Annual review of applied linguistics, 18*, 81-101.
- Mueller, G. A. (1980). Visual contextual cues and listening comprehension: An experiment. *The Modern Language Journal, 64(3)*, 335-340.
- Nation, I. S. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review 63*, 59-82.
- Nunan, D. (1999). *Second Language Teaching & Learning*. Boston, MA: Heinle & Heinle Publishers.

- O'Brien, L. (1995). The Learning Channel preference checklist. J. M. Reid içinde, *Learnin Styles in the ESL/EFL Claroom* (s. 196-201). Boston, MA: Heinle & Heinle.
- Ockey, G. J. (2007). Construct implications of including still image or video in computer-based listening tests . *Language Testing*, 24(4), 517-537.
- Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, 3, 53-66.
- Oxford, R. L. (1993). Research update on teaching L2 listening. *System*, 21(2), 205-211.
- Oxford, R. L. (1995). Style analysis survey (SAS): Assessing your own learning and working style. J. M. Reid içinde, *Learning styles in the ESL/EFL classroom* (s. 208- 215). Boston, MA: Heinle & Heinle.
- Pallant, J. (2001). *SPSS survival manual* . Maidenhead, PA: Open University Press.
- Park, G. (2004). Comparison of L2 listening and reading comprehension by university students learning English in Korea. . *Foreign Language Annals*, 37(3) , 448-458.
- Parry, T. S., & Meredith, R. A. (1984). Videotape vs. audiotape for listening comprehension tests: An experiment. *OMLTA journal* , 47-53.
- Pearce, C. G., Johnson, I. W., & Barker, R. T. (2003). Assessment of the listening styles inventory progress in establishing reliability and validity. *Journal of Business and Technical Communication*, 17(1), 84-113.

- Progosh, D. (1996). Using video for listening assessment opinions of test-takers. *TESL Canada Journal* 14 (1), 34-44.
- Purdy, M. (1997). What is listening? In M. Purdy, & D. Borisoff, *Listening in everyday life: A personal and professional approach (2nd ed.)* (pp. 1-20). Lanham, MD: University Press of America.
- Read, J. (2002). The use of interactive input in EAP listening assessment. *Journal of English for Academic Purposes* 1, 105-119.
- Reid, J. (1995). Perceptual learning style preference questionnaire. J. M. Reid içinde, *Learning styles in the ESL/EFL classroom* (s. 202-207). Boston, MA: Heinle & Heinle.
- Richards, J. C. (2006). Materials development and research—making the connection. *RELC Journal*, 37(1), 5-26.
- Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3-13.
- Rost, M. (1990). *Listening in Language Learning*. New York: Longman.
- Rost, M. (2011). *Teaching and researching listening (2nd ed.)*. Harlow, UK: Pearson.
- Rubin, D. L. (1993). Listenability=oral-based discourse + considerateness. In A. D. Wolvin, & C. G. Coakley (Eds), *Perspectives* (pp. 261-281). Norwood, NJ: Ablex Publishing.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78, 199-221.

- Rubin, J. (1995). The contribution of video to the development of competence in listening. In D. Mendelsohn, & J. Rubin, *A guide for the teaching of second language listening* (pp. 151-165). San Diego, CA: Dominie Press.
- Schmidt-Rinehart, B. (1994). The effects of topic familiarity on second language listening comprehension. . *The modern language journal*, 78(2), 179-189.
- Scholz, G. E. (1993, April). Exploring ESP and language testing. Paper presented at the annual meeting of the Southeast Asian Ministers of Education Organization Regional Language Center Seminar. Singapore.
- Schriver, K. A. (1997). *Dynamics in document design: creating text for readers*. New York, NY: John Wiley & Sons, Inc.
- Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *The Modern Language Journal*, 76(4), 480-490.
- Shin, D. (1998). Using video-taped lectures for testing academic language. *International Journal of Listening*, 12, 56-79.
- Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language testing*, 8(1), 23-40.
- SPSS. (2009). *PASW Statistics for Windows, Version 18.0*. Chicago: SPSS Inc.
- Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31, 577-607.
- Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55(4), 661-699.

- Suvorov, R. (2009). Context visuals in L2 listening tests: The effect of photographs and video vs. audio-only format. *Developing and evaluating language learning materials*, 53-68.
- Suvorov, R. (2013). Interacting with visuals in L2 listening tests: An eye-tracking study (Doctoral thesis). Ames, IA: Iowa State University.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review*, 22(2), 123-138.
- Tarrant, W., Ware, J., & Mohammed, A. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Medical Education*, 9(1), 1-8.
- Tauroza, S., & Luk, J. (1997). Accent and second language listening comprehension. *RELC Journal*, 28(1), 54-71.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education*, 2, 53.
- Taylor, L., & Geranpayeh, A. (2011). Assessing listening for academic purposes: defining and operationalising the test construct. *Journal of English for Academic Purposes*, 10, 89-101.
- Tsui, A. B., & Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics* 19, 432-51.
- Türkyılmaz, M. (2010). Bir ders aracı olarak dinleme metinlerinin anlamaya etkisi: "eskici" metni örneğinde deneysel bir çalışma. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 23(1), 207-220.

- Tyler, M. (2001). Resource consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning*, 51(2), 257-280.
- Underwood, M. (1989). *Teaching listening*. New York: Longman.
- Ur, P. (2009). *Teaching Listening Comprehension*. Cambridge: Cambridge University Press.
- Vandergrift, L. (2003). 'Orchestrating strategy use: Toward a model of the skilled second language listener'. *Language Learning*, 53(3), 463-96.
- Vandergrift, L. (2004). 1. Listening to Learn or Learning to Listen? *Annual Review of Applied Linguistics*, 24, 3-25.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40, 191-210.
- Vandergrift, L., & Goh, C. C. (2012). *Teaching and learning second language listening: Metacognition in action*. New York: Routledge.
- Wagner, E. (2007). Are they watching? Test-taker viewing behavior during an L2 video listening test. *Language Learning & Technology*, 11(1), 67-86.
- Wagner, E. (2008). Video listening tests. What are they measuring? *Language Assessment Quarterly*, 5(3), 218-243.
- Wagner, E. (2010a). Test-takers' interaction with an L2 video listening test. *System*, 38(2), 280-291.

- Wagner, E. (2010b). The effect of the use of video texts on ESL listening test-taker performance. *Language Testing*, 27(4), 493-513.
- Wagner, E. (2013). An investigation of how the channel of input and access to test questions affect L2 listening test performance. *Language Assessment Quarterly*, 10(2), 178-195.
- Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching* 31, 57-71.
- Wilberschied, L., & Berman, P. M. (2004). Effect of using photos from authentic video as advance organizers on listening comprehension in an FLES Chinese class. *Foreign Language Annals*, 37(4), 534-540.
- Yanagawa, K., & Green, A. (2008). To show or not to show: The effects of item stems and answer options on performance on a multiple-choice listening comprehension test. *System*, 36(1), 107-122.
- Zhao, Y. (1997). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics*, 18(1), 49-68.
- Zwiers, J. (2002). Teacher practices and perspectives for developing academic language. *International Journal of Applied Linguistics*, 17(1), 93-116.

APPENDICES

Appendix A. Demographic Information and Information of English Listening Questionnaire (in English)

QUESTIONNAIRE – DEMOGRAPHIC INFORMATION AND INFORMATION OF ENGLISH LISTENING

Directions: Please circle the item that best fits you or write a short answer in each item.

1. Gender: **Male** **Female**

2. Age: **less than 18** **18 – 20** **21 - 23** **more than 23 years
old**

3. Department:

a. Psychology

f. Economics

b. Sociology

g. Political Science and Public

c. Turkish Language and Literature

Administration

d. International Relations

h. International Trade and Finance

e. Business Administration

4. What is your English level in the preparatory school?

A2

B1

B2

5. Total length of studying English (formal instruction given in schools or institutions):

a. less than 1 year

b. 1 - 4 years

c. 5 - 10 years

d. more than 10 years

6. Do you watch English TV programs or movies? (If your answer is “No”, please go to **question #9**)

a. Yes

b. No

7. If you answer “Yes” to **question #6**, how much time do you spend watching English TV programs or movies?

a. Less than 30 minutes a day

b. 30 min.-1 hour a day

- c. More than 1 hour a day
 - d. Other _____
8. How much do you understand when you watch English programs or movies?
- a. Less than 25%
 - b. 25% - 50%
 - c. 51% - 75%
 - d. More than 75%
9. Do you listen to English radio? (If your answer is “No”, please go to **question #12**)
- a. Yes b. No**
10. If you answer “Yes” to **question #9**, how much time do you spend listening English radio?
- a. Less than 30 minutes a day
 - b. 30 min.-1 hour a day
 - c. More than 1 hour a day
 - d. Other _____
11. How much do you understand when you listen to English radio?
- a. Less than 25%
 - b. 25% - 50%
 - c. 51% - 75%
 - d. More than 75%
12. Which way do you think is the best for improving your English listening skill?
- a. Watching TV programs or movies (not for entertainment)
 - b. Listening to the radio (not for entertainment)
 - c. Listening podcasts (lectures, courses, seminars) on the Internet
 - d. Listening to English lessons given by a teacher in the classroom
 - e. Listening to English native friends
 - f. Other _____
13. How much do you understand when you listen to lessons given in English?
- a. Less than 25%
 - b. 25% - 50%
 - c. 51% - 75%
 - d. More than 75%

**Appendix B. Demographic Information and Information of English Listening
Questionnaire (in Turkish)**

İNGİLİZCE DİLİNDE DİNLEME VE DEMOGRAFİK BİLGİ ARAŞTIRMASI

Yönerge: Aşağıdaki sorulara size uyan en uygun cevabı lütfen ya halka içine alın ya da kısa cevaplar verin.

1. Cinsiyet: **Bay** **Bayan**

2. Yaş: **18 yaş altı** **18 – 20** **21 - 23** **23 yaş üstü**

3. Bölümünüz:

a. **Psikoloji**

e. **İşletme**

b. **Sosyoloji**

f. **İktisat**

c. **Türk Dili ve Edebiyatı**

g. **Siyaset Bilimi ve Kamu Yönetimi**

d. **Uluslararası İlişkiler**

h. **Uluslararası Ticaret ve Finans**

4. Hazırlık okulundaki seviyeniz nedir?

A2

B1

B2

5. Bugüne kadar ki toplam İngilizce öğrenim süreniz (okul veya kurslarda alınan eğitim)

a. **1 yıldan az**

b. **1 - 4 yıl arası**

c. **5 - 10 yıl arası**

d. **10 yıldan fazla**

6. İngilizce televizyon programı ya da film izler misiniz? (Cevabınız Hayır ise **9. Soruya geçin**)

a. **Evet**

b. **Hayır**

7. Eğer **6. Soruya** cevabınız “Evet “ ise, bu İngilizce programları veya filmleri ne sıklıkla izlersiniz?

a. **Günde 30 dakikadan az**

b. **Günde 30 dakika ile 1 saat arası**

c. **Günde 1 saatten fazla**

- d. Başka belirtmek istediğiniz seçenek _____
8. İngilizce bir program veya film izlediğinizde bunun ne kadarını anlayabiliyorsunuz?
- 25% in den az
 - 25%- 50% arası
 - 51% - 75% arası
 - 75% in den fazla
9. İngilizce radyo programı dinler misiniz? (Cevabınız Hayır ise **12. Soruya** geçin)
- Evet
 - Hayır
10. Eğer 9. Soruya cevabınız “Evet “ ise, bu programları ne sıklıkla dinlersiniz?
- Günde 30 dakikadan az
 - Günde 30 dakika ile 1 saat arası
 - Günde 1 saatten fazla
 - Başka belirtmek istediğiniz seçenek _____
11. İngilizce bir radyo programı dinlediğinizde bunun ne kadarını anlayabiliyorsunuz?
- 25% in den az
 - 25%- 50% arası
 - 51% - 75% arası
 - 75% in den fazla
12. Sizce **İngilizce dinleme becerinizi** en iyi geliştiren yöntem hangisidir?
- İngilizce televizyon programı ya da film izlemek (eğlence amaçlı değil)
 - İngilizce radyo programı dinlemek (eğlence amaçlı değil)
 - İnternet üzerinden İngilizce podcast (seminer, ders, akademik konuşma) dinlemek
 - Sınıfta öğretmenin anlattığı İngilizce dersini dinlemek
 - Ana dili İngilizce olan bir arkadaşımızı dinlemek.
 - Başka belirtmek istediğiniz _____
13. Sınıfta İngilizce olarak anlatılan dersin ne kadarını anlayabiliyorsunuz?
- 25% in den az
 - 25%- 50% arası
 - 51% - 75% arası
 - 75% in den fazla

Appendix C. A Letter for Expert Opinion of EFL Listening Scale (in Turkish)

Sayın Hocam;

“Yabancı dil olarak İngilizce öğrenen üniversite hazırlık sınıfı öğrencilerin, akademik dinleme aktivitelerinde kullanılan ortam ve içerik görsellerinin dinlenen konuyu anlamaya olan etkisini ve öğrencilerin bireysel farklılıklarının anlamaya olan etkilerini” araştıran doktora tezinde kullanılmak üzere bir ölçek geliştirilmektedir. Ölçekle, yabancı dil olarak İngilizce öğrenen öğrencilerin akademik dinleme materyallerinde tercih ettikleri özellikleri belirlenmeye çalışılmaktadır. Ölçek **Ortam ve Görseller boyutlarından** oluşmaktadır. Ölçekteki ölçütlerin, amaçlanan ortam için önem derecesi, anlaşılabilirliği ve ilgili alt boyutu temsil etme derecesi hakkında sizlerin görüşlerine gereksinim duyulmaktadır. Bu çerçevede aşağıdaki derecelendirmeyi esas alarak ölçeği değerlendirmeniz, çalışmanın gerçekleşmesi ve anlamlı sonuçlar vermesi açısından son derece önemlidir. Ayıracağınız emek ve zaman için şimdiden teşekkür ediyoruz.

Danışman: Yard. Doc. Dr. Gonca Kızılkaya Cumaoğlu

Doktora Öğrencisi: Hümeysra Genç

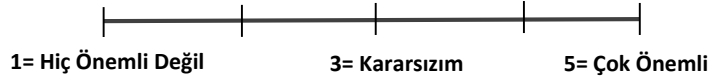
Boyutların Açıklaması

Ortam: Dinleme aktivitesinin yapıldığı ortam, bireysel ve toplu yapılan dinleme ile ilgili olası ilkeleri içermektedir.

Görseller: Dinleme aktivitesi sırasında, anlamayı kolaylaştırmak için öğrencinin hangi tür görselleri (ortam ya da içerik) tercih ettiği, ya da hiç tercih etmediği ile ilgili olası ilkeleri içermektedir.

Değerlendirmede Esas Alınan Derecelendirme

Ölçütün Önem Derecesi:



İfadenin Anlaşılabilirlik Derecesi:



İlgili alt boyutla ilişkililik Derecesi:



Öğrenim Durumunuz	
Alanınız	
Yaşınız	
Mesleğiniz	

YABANCI DİL OLARAK İNGİLİZCE ÖĞRENEN ÖĞRENCİLERİN DİNLEME MATERYALLERİNDE TERCİH ETTİKLERİ ÖZELLİKLER		Önemli bir ölçüt	Anlaşılır	İlgili alt boyutla ilişkili	Öneri
Ortam	1. Tek başıma dinleme yapmayı, sınıfta toplu halde dinleme yapmaya tercih ederim.				
	2. Dinleme aktivitesini yaptığım fiziksel ortam dinlediğim parçayı anlayabilmem için önemlidir.				
	3. Bilgisayar ekranı karşısında kulaklıkla dinleme yapmayı tercih ederim.				
	4. Amfi gibi geniş ortamlarda dinleme başarımlarım düşüyor.				
Görseller	5. Dinleme parçasının konuyla ilgili fotoğraf ya da resimlerle desteklenmesi anlamımı kolaylaştırır.				
	6. Dinleme parçasının veri tablolarıyla desteklenmesi anlamımı kolaylaştırır.				

7. Dinleme sırasında konuşmanın yapıldığı ortamı görmek (sınıf, restoran, ofis gibi) dinlediğim parçayı anlamamı kolaylaştırır.				
8. Dinleme sırasında konuşmacıyı görmek dinlediğim parçayı anlamamı kolaylaştırır.				
9. Dinleme sırasında konuşmacıyı görmesem bile parçayı kolaylıkla anlarım.				
10. Görüntülü dinleme materyellerinde gördüğüm görüntüden ziyade dinlediğim daha çok aklımda kalır.				
11. Görüntülü dinlemelerde konuşmacının el-kol hareketleri, mimikleri konuyu anlamamda yardımcı olur.				
12. Dinleme sırasında parçanın konusunu hayalimde canlandırırım.				
13. Dinleme parçasındaki konuşmacıları hayalimde canlandırırım.				
14. Dinleme sırasında konuşmanın geçtiği ortamı hayalimde canlandırırım.				
15. Dinleme sırasında anladığım anahtar kelimeleri (örneğin telefon numarası, bir kişinin fiziksel özellikleri) aklımda hemen resmederek canlandırırım.				
16. Görsel materyallerle desteklenmiş dinleme parçalarını desteklenmemişlere tercih ederim.				
17. Video dinlemelerini, görüntüsüz (radyo/telefon gibi sadece ses olarak) dinlemelere tercih ederim.				
18. Video dinlemelerinde görüntüye dalıp farkında olmadan dinlemeyi bıraktığım oluyor.				

19. Video dinlemelerinde konuyla ilgili görüntüler konuyu anlamamda yardımcı oluyor.				
20. Dinleme yaparken kâğıda konuyla ilgili şekiller çizerim.				
21. Dinleme sırasında kimin söylediğinden ziyade ne söylediğine önem veririm.				
22. Dinleme sırasında konuşmacının aksanı konuşulan konuyu anlamam için önemlidir.				
23. Dinleme sırasında konuşmacının ses tonu konuşulan konuyu anlamam için önemlidir.				
24. Dinleme sırasında konuşmacının vurgulamaları konuşulan konuyu anlamam için önemlidir.				
25. Dinleme becerisini, okuma, yazma, becerilerine kıyasla daha zor olduğunu düşünüyorum.				

Appendix D. English as Foreign Language Listening Style Scale (in Turkish)

YABANCI DİL OLARAK İNGİLİZCE DİNLEME STİLLERİ ÖLÇEĞİ

Aşağıdaki ölçekte yabancı dil (İngilizce) dinleme etkinliklerinde ki stilleriniz ile ilgili çeşitli önermeler verilmiştir. Önermelere katılma derecenizi size uygun gelen kutucuktaki rakamı işaretleyerek belirtiniz. Verdiğiniz yüksek puan, o önermeye ne kadar çok katıldığınızı, düşük puan ne derecede katılmadığınızı gösterir. Doğru ya da yanlış cevap yoktur. Lütfen boş bırakmayınız. İsminizi yazmayınız.

Önermelere Katılım Derecelendirmesi:

- 1: Kesinlikle Katılmıyorum
- 2: Katılmıyorum
- 3: Kararsızım
- 4: Katılıyorum
- 5: Kesinlikle Katılıyorum

Yaşınız: _____ Kurunuz: A2 B1 B2

Bölümünüz: _____ Cinsiyetiniz: Kız Erkek

		1 Kesinlikle Katılmıyorum	2 Katılmıyorum	3 Kararsızım	4 Katılıyorum	5 Kesinlikle Katılıyorum
1	Dinleme parçasının konuyla ilgili fotoğraf ya da resimlerle desteklenmesi anlamamı kolaylaştırır.	1	2	3	4	5
2	Dinleme sırasında konuşmanın yapıldığı ortamı görmek (sınıf, restoran, ofis gibi) dinlediğim parçayı anlamamı kolaylaştırır.	1	2	3	4	5
3	Görüntülü dinlemelerde konuşmacının el-kol hareketleri, mimikleri konuyu anlamamda yardımcı olur.	1	2	3	4	5
4	Dinleme sırasında konuşmanın geçtiği ortamı hayalimde canlandırırım.	1	2	3	4	5
5	Dinleme sırasında konuşmacının vurgulamaları konuşulan konuyu anlamam için önemlidir.	1	2	3	4	5
6	Dinleme sırasında konuşmacının telaffuzu anlamamı zorlaştırırsa dinlemeyi bırakırım.	1	2	3	4	5

		1 Kesinlikle Katılmıyorum	2 Katılmıyorum	3 Kararsızım	4 Katılıyorum	5 Kesinlikle Katılıyorum
7	Dinleme parçasının veri tablolarıyla desteklenmesi anlamımı kolaylaştırır.	1	2	3	4	5
8	Dinleme parçasındaki konuşmacıları hayalimde canlandırırım.	1	2	3	4	5
9	Dinleme sırasında konuşmacıyı görmek dinlediğim parçayı anlamamı kolaylaştırır.	1	2	3	4	5
10	Dinleme sırasında konuşmacının aksanı konuşulan konuyu anlamam için önemlidir.	1	2	3	4	5
11	Dinleme parçasındaki anahtar kelimeleri anlarsam parçanın tamamını anlarım.	1	2	3	4	5
12	Görsel materyaller içeren dinleme parçaları içermeyenlere göre daha çok aklımda kalır.	1	2	3	4	5
13	Dinleme sırasında konuşmacının ses tonu konuşulan konuyu anlamam için önemlidir.	1	2	3	4	5
14	Dinleme sırasında parçanın konusunu hayalimde canlandırırım.	1	2	3	4	5
15	Görüntülü dinleme materyellerinde gördüğüm görüntüden ziyade dinlediğim daha çok aklımda kalır.	1	2	3	4	5
16	Dinleme sırasında parçanın ana fikrinden çok, öncelikle detaylara odaklanırım.	1	2	3	4	5
17	Görüntülü dinlemelerde konuyla ilgili görüntüler konuyu anlamama yardımcı olur.	1	2	3	4	5
18	Dinleme parçasını anlayabilmem için parçadaki bütün kelimeleri anlamam önemlidir.	1	2	3	4	5
19	Dinleme sırasında anladığım anahtar kelimeleri (örneğin telefon numarası, bir kişinin fiziksel özellikleri) aklımda hemen resmederek canlandırırım.	1	2	3	4	5

Appendix E. The Last Version of English as Foreign Language Listening Style Scale (in English)

ENGLISH AS A FOREIGN LANGUAGE LISTENING STYLE SCALE

This scale has been designed to help you identify the ways you prefer to listen in English language. In the below scale various statements about your listening styles in English listening activities are given. Please read each statement in the scale. Then circle the number in the box that best represents the way you prefer while listening in the English language. **The highest number shows how much you agree with each statement and the lowest number shows how much you disagree with each statement.** There is no right or wrong response.

Please respond to all statements.

Rating Scale for the Statements:

- 1: Strongly Disagree
- 2: Disagree
- 3: Undecided
- 4: Agree
- 5: Strongly Agree

		1 Strongly Disagree	2 Disagree	3 Undecided	4 Agree	5 Strongly Agree
1	That the listening text is supported by photos or pictures related to topic makes it easy to understand	1	2	3	4	5
2	While listening, observing the environment (such as classroom, restaurant or office) where the conversation is carried out makes the listening activity easy to understand.	1	2	3	4	5
3	In video listening, the gestures and the mimics of the speaker help me understand the topic.	1	2	3	4	5
4	I visualize the location where the conversation takes place in my head while listening.	1	2	3	4	5
5	The emphasis that the speaker makes during listening is important to me to understand the spoken topic.	1	2	3	4	5
6	That listening text are supported by data tables helps my understanding.	1	2	3	4	5
7	I visualize the speakers who are talking in the listening text.	1	2	3	4	5
8	Seeing the speaker while listening helps my understanding the listening text better.	1	2	3	4	5
		1	2 Disagree	3 Undecided	4 Agree	5

		Strongly Disagree				Strongly Agree
9	The accent of the speaker during listening is important to me to understand the spoken topic.	1	2	3	4	5
10	If I understand the key words in the listening text, I can understand all listening parts.	1	2	3	4	5
11	Listening texts including visual materials remains longer in my mind than the ones not including visual materials.	1	2	3	4	5
12	The speaker's tone of voice during listening is important to me to understand the spoken topic.	1	2	3	4	5
13	While listening I visualize the topic of the listening text in my head.	1	2	3	4	5
14	While listening, I primarily focus on the details rather than the main idea of the listening text.	1	2	3	4	5
15	The visuals related to the topic in video listening activities help me understand the topic.	1	2	3	4	5
16	It is important for me to understand all words spoken in the listening text in order to understand the listening part.	1	2	3	4	5
17	I visualize the key words I catch while listening (such as phone numbers or physical features) by picturing them in my mind immediately.	1	2	3	4	5

Appendix F. The Last Version of English as Foreign Language Listening Style Scale (in Turkish)

YABANCI DİL OLARAK İNGİLİZCE DİNLEME STİLLERİ ÖLÇEĞİ

Bu ölçek İngilizce dilinde dinleme yaparken tercih ettiğiniz yolları belirlemeye yardımcı olmak için oluşturulmuştur. Aşağıdaki ölçekte İngilizce dinleme etkinliklerindeki dinleme stilleriniz ile ilgili çeşitli önermeler verilmiştir. Ölçekteki her bir önermeyi lütfen okuyunuz. Sonra, İngilizce dilinde dinleme yaparken tercih ettiğiniz yolları en uygun ifade eden kutucuktaki rakamı işaretleyiniz. **Yüksek puan, o önermeye ne kadar çok katıldığınızı, düşük puan ne derecede katılmadığınızı gösterir.** Doğru ya da yanlış cevap yoktur. **Lütfen bütün önermelere cevap veriniz.**

Önermeler için Derecelendirme Ölçeği:

- 1: Kesinlikle Katılmıyorum
- 2: Katılmıyorum
- 3: Kararsızım
- 4: Katılıyorum
- 5: Kesinlikle Katılıyorum

		1 Kesinlikle Katılmıyorum	2 Katılmıyorum	3 Kararsızım	4 Katılıyorum	5 Kesinlikle Katılıyorum
1	Dinleme parçasının konuyla ilgili fotoğraf ya da resimlerle desteklenmesi anlamamı kolaylaştırır.	1	2	3	4	5
2	Dinleme sırasında konuşmanın yapıldığı ortamı görmek (sınıf, restoran, ofis gibi) dinlediğim parçayı anlamamı kolaylaştırır.	1	2	3	4	5
3	Görüntülü dinlemelerde konuşmacının el-kol hareketleri, mimikleri konuyu anlamamda yardımcı olur.	1	2	3	4	5
4	Dinleme sırasında konuşmanın geçtiği ortamı hayalimde canlandırırım.	1	2	3	4	5
5	Dinleme sırasında konuşmacının vurgulamaları konuşulan konuyu anlamam için önemlidir.	1	2	3	4	5
6	Dinleme parçasının veri tablolarıyla desteklenmesi anlamamı kolaylaştırır.	1	2	3	4	5
7	Dinleme parçasındaki konuşmacıları hayalimde canlandırırım.	1	2	3	4	5
8	Dinleme sırasında konuşmacıyı görmek dinlediğim parçayı anlamamı kolaylaştırır.	1	2	3	4	5
		1	2 Katılmıyorum	3 Kararsızım	4 Katılıyorum	5

		Kesinlikle Katılmıyorum				Kesinlikle Katılıyorum
9	Dinleme sırasında konuşmacının aksanı konuşulan konuyu anlamam için önemlidir.	1	2	3	4	5
10	Dinleme parçasındaki anahtar kelimeleri anlarsam parçanın tamamını anlarım.	1	2	3	4	5
11	Görsel materyaller içeren dinleme parçaları içermeyenlere göre daha çok aklımda kalır.	1	2	3	4	5
12	Dinleme sırasında konuşmacının ses tonu konuşulan konuyu anlamam için önemlidir.	1	2	3	4	5
13	Dinleme sırasında parçanın konusunu hayalimde canlandırırım.	1	2	3	4	5
14	Dinleme sırasında parçanın ana fikrinden çok, öncelikle detaylara odaklanırım.	1	2	3	4	5
15	Görüntülü dinlemelerde konuyla ilgili görüntüler konuyu anlamama yardımcı olur.	1	2	3	4	5
16	Dinleme parçasını anlayabilmem için parçadaki bütün kelimeleri anlamam önemlidir.	1	2	3	4	5
17	Dinleme sırasında anladığım anahtar kelimeleri (örneğin telefon numarası, bir kişinin fiziksel özellikleri) aklımda hemen resmederek canlandırırım.	1	2	3	4	5

Appendix G. The Audio script of Part-1 (Context Audio) in Listening Test-1

People who made a great impact on our society

Every year different organizations give awards for global citizenship or personal social responsibility. Some who receive awards are famous, such as Mother Theresa and Bill Gates. Others come from the growing number of ordinary people who have made an impact on society on a smaller scale. Our topic today is extraordinary work for causes by people who might not be in the spotlight.

First, I present Derreck Kayongo, a Ugandan refugee. When Kayongo was working for CARE International, he thought about the children in Africa who were dying. The cause was not war or famine alone. Unsanitary conditions created a problem that he could not ignore. The solution was as simple as soap and water. More than two million children in poor countries die each year from diseases that could be prevented by better hygiene. Yet hotels in developed countries throw away millions of partially used soap bars every day. In 2009, Kayongo started the Global Soap Project. The GSP volunteers collect the soap from hotels. The organization then sanitizes and recycles them. As of 2012, GSP has distributed over 250,000 bars to 21 countries including Haiti, South Sudan, Kenya, Honduras, and Afghanistan.

Second, I would like to present John Francis' unusual story. When he was a young man in California, he witnessed something that would change his life forever. Two oil tankers had collided in the San Francisco bay. The resulting oil spill killed birds and sea life. Francis realized that he was partly responsible for the spill because he was one of millions of people using gas vehicles. In 1970, he decided that he would walk everywhere and stop riding in cars or trucks. Shortly after, he also took a

vow of silence. Francis walked all over California. Then he walked from one end of the United States to the other on behalf of the environment.

How did society benefit from his actions? Francis studied for a PhD in environmental studies. After ending his silence in 1990, he became a goodwill ambassador for the United Nations Environmental Program. Later he contributed to the U.S. Coast Guard's Oil Pollution Act of 1990. He also founded Planetwalk, a nonprofit environmental education organization

That's all we have time for today. In the next class I will read stories of a few more unknown heroes who have made outstanding contributions to society. Then you will decide who we would recognize with our own award.

Source: taken from the exam materials of Q Skills for Success Listening and Speaking

Appendix H. The Audio script of Part-2 (Content Audio) in Listening Test-1

McDonald`s

Fast food is a term for food made quickly in a restaurant and served in a package. Everyone knows that McDonald's is the world's largest fast-food chain. According to the company website, there are more than 34,000 restaurants in 118 countries. Those restaurants serve almost 70 million customers every day. In *Fast Food Nation* a book about fast-food in America, the author explains that when McDonald's opened in 1954, it was specifically aimed at children. In 1979, McDonald's started offering the Happy Meal. Nutritionists, however, are not happy with a fast-food nation. They blame restaurants like McDonald's for an increase in obesity. McDonald's response is that they are only providing what the customer wants. Unfortunately, it appears that the trend is to choose taste over nutrition.

In 2008, New York was the first city to require fast-food restaurants to put calorie counts on their menus. Now the US Food and Drug Administration wants nutrition information on all menus in chain restaurants. Could it really be so easy? When consumers see how many calories they eat, will they be wise enough to make the healthy choice? Well, the answer has surprised many people from nutritionists to the restaurant owners themselves. It seems that, at least in fast-food restaurants, the choice consumers make is not always the best one.

What's the problem? A professor of health policy at NYU found that 54 percent of a survey group did notice the calories written on the menus. However, only one quarter of those customers said they then consumed fewer calories as a result. One food consultant has pointed out that people usually don't think of fast-food restaurants as a healthy place to eat. As a result, they just don't pay attention.

Still, restaurants like McDonald's have made an effort. What do you think happened when the restaurant started spending 1/6 of its advertising time on salads? Did customers realize they could get healthy food at McDonald's? Did they choose the salads instead? Nope. The salads made up only 2 to 3 percent of sales. Maybe it's just the consumers' fault. Another professor, this one from Duke University, also found that customers often made the wrong choice. When the menu included a mix of unhealthy, neutral, and healthy items, seeing the healthy foods was enough to make the customers feel good. Then they went ahead and made the unhealthy choice more often.

Obviously, fast-food restaurants like McDonald's are in business to make money. If customers want an unhealthy diet, the restaurants are ready to feed them.

Source: taken from the exam materials of Q Skills for Success Listening and Speaking

Appendix I. The Audio script of Part-3 (Context Video) in Listening Test-1

Cigarette Smoking

Hello, in this film I'm going to talk to you about two of the biggest problems facing us globally in public health. The first is the epidemic of tobacco which in historical terms is a relatively new problem and the second is tuberculosis which in historical terms is a much older problem.

Tobacco smoking is perhaps one of the most puzzling of human behaviors. You might recall, there's a very, very funny American comedian called Bob Newhart who plays on the absurdities of cigarette smoking in an imaginary conversation with Sir Walter Raleigh, when they discovered tobacco in the, in, in the, in the American colonies. And Raleigh is trying to explain to him how good it is to crush up leaves, put them in a piece of paper, set fire to them, and put them in your mouth. And of course this is met with incredulity and the audience thinks it's hilarious.

But there's a serious under, underlying point to this, because the thing that drives cigarette smoking of course is not the smoke itself, it's the nicotine that's in the tobacco. Nicotine is about one of the most highly addictive substances known to mankind, and it's very, very difficult to quit. Nicotine itself is, other things considered, a relatively harmless substance. It's the toxins in the tobacco smoke that cause the harm. And my word, don't they cause harm? Everything from lung cancer, emphasize, bronchitis, heart disease, hypertension, and a range of other problems that can be laid firmly at the door of cigarette smoking.

It only became fashionable really towards the end of the 19th century first among the rich, then among rich men, then among all men and then the population as

a whole. Reaching its height in the western world in the 1950s followed swiftly by epidemics of lung cancer, heart disease, and related problems.

In global terms, what we see today, a rather different prevalence of smoking around the world. In countries like the United States, United Kingdom, it's about 20%. But in China it's around 75% among men. And so we see great variation, and of course, the epidemics that follow in the, in the wake of cigarette smoking will follow these trends, too.

But what can be done about it given the powerful addiction? The first real scientific evidence which pointed to the dangers associated to cigarette smoking. With cigarette smoking, is a paper published in 1952 by Dalton Hill to British scientists who identified the link between cigarette smoking and lung cancer. Ten years later, Royal College of Physicians in London brought out a definitive study showing unequivocally the links between poor health and cigarette smoking.

Two years later in the United States the Surgeon General, an even larger array of evidence nailing the problem of cigarette smoking as the major problem of mid20th century the mid20th century. Since then, there have been great strides forward.

I think one can argue that in some countries at least the war against tobacco has been one of the most successful public health triumphs of the modern age. So we see the rates of smoking falling to the 20% or so that I mentioned a moment.ago.

Source: available at www.coursera.org

Appendix J. The Audio script of Part-4 (Content Video) in Listening Test-1

Global Warming

For 2.5 million years ago, The Earth's climate was fluctuated. Cycling from ice ages to warmer periods. But in the last century, the planet's temperature has risen unusually fast about 1.2 to 1.4 degrees Fahrenheit. Scientist believe it's human activity that's driven the temperatures up. A process known as global warming. Ever since the industrial revolution began, factories, power plants and eventually cars have burned fossil fuels such as oil and coal releasing huge amount of carbon dioxide and other gases into the atmosphere. These greenhouse gasses trap heat near the Earth through a naturally occurring process called the greenhouse effect.

The greenhouse effect begins with the sun and the energy that radiates to the Earth. The earth and the atmosphere absorb some of this energy while rest is radiated back into space. Naturally occurring gasses in the atmosphere trap some of this energy and reflected back warming the earth. Scientists now believe that the greenhouse effect is being intensified by the extra greenhouse gasses that humans have released.

Evidence for global warming includes a recent string of very warm years. Scientists report that 1998 was the warmest year in measured history. With 2005 coming in the second. Meanwhile, radiance taken from ice cores show that the greenhouse gasses, carbon dioxide and methane have hit their highest level in the past 420,000 years. Arctic sea ice is also shrinking. According to NASA studies, the extent of arctic sea ice has declined 10% in the last 30 years. As long as industrialized nations consume energy and developing countries increase their fossil fuel consumption, the concentration of greenhouse gasses in the atmosphere will continue to rise.

Researchers predict that temperatures will increase about 2 to 10 degrees Fahrenheit by the end of the century. What is less certain is what rising temperature's mean for the planet. Some climate models predict subtle changes. Others forecast rising sea levels which could flood coastal areas around the world. Weather patterns could change making hurricanes more frequent. Severe drought could become more common in warm areas and species unable to adapt to the changing conditions would face extinction.

Although much remains to be learned about global warming, many organizations advocate cutting greenhouse gas emissions to reduce the impact of global warming. Consumers can help by saving energy around the house switching to compact fluorescent light bulbs and driving fewer miles in the car. Retreating these simple changes may help keep the earth cooler in the future.

Source: available at <https://www.youtube.com/watch?v=oJAbATJCugs>

Appendix K. The Audio script of Part-5 (Audio-Only) in Listening Test-1

First Impression of Food

In class today, I'll be discussing how we form our impressions of food. First, let's review the five senses that contribute to our perceptions of everything around us. Remember, those senses are sight, smell, hearing, touch, and taste. Immediately, most of us think of taste as the primary sense that affects our impression of foods and whether or not we like them. However, think about the information we get before we even put food in our mouths. It may be that other senses are more reliable.

Let's start with some typical experiences. Imagine I'm a student home for vacation. I've just finished my final exams and wake up anxious for a homemade breakfast. Probably while I'm lying in bed, I smell the coffee. The smell actually helps me get out of bed. Perhaps I heard my father grinding the coffee beans first. I assume the coffee will be better than the instant coffee I make in my dorm room. So here the smell has influenced my impression of the coffee even before I taste it.

Do you think you can form a negative impression of a food from hearing its name? Well, ask yourself if your first impression of a fish called 'spotted weakfish' or 'bigmouth sleeper' is a positive one. They do not sound very appetizing, do they? That's one way hearing might influence the foods we select. And of course just the sound of food cooking, such as a sizzling stir-fry, can make our mouths water.

Touch and sight also play a role. While some people love sushi, others refuse to eat it because they don't like the way it feels in their mouths. Some eaters say that a lot of food piled on a plate makes them lose their appetite. Others may be suspicious of a bowl of meat covered with a sauce they are not familiar with. Now

many menus show pictures of food. Chefs are worried about the presentation. If the plate looks beautiful, we believe that the food will be delicious.

Now let's get back to the sense that we all thought was the most important, taste. You can see that many other senses have come into play before we actually taste the food. The food may taste better or worse than we thought based on the way it sounded, looked, felt, or smelled. Did we make a snap judgment? So maybe taste is *not* the most important sense in creating our first impressions. It just confirms them. And of course, taste may influence our future choices - we may decide we love Thai food because it is spicy or we hate English food because it isn't. I'll end with an interesting example. The durian, an Asian fruit, has a sharp pointy exterior and an awful smell. However, many are able to get past their first impressions and enjoy the taste of this unusual fruit.

Source: taken from the exam materials of Q Skills for Success Listening and Speaking

Appendix L. The Audio script of Part-1 (Audio-Only) in Listening Test-2

Greeting Cards

My boss was going into the hospital for surgery last week. I thought about sending an email to wish her well. Would that be too much like a work memo? I thought about sending her an e-card. Would she have her computer at the hospital? A phone call seemed too personal. Then, out of the blue, an old-fashioned idea came to mind. Why not stop at a local store and pick up a greeting card? I worried that there might not actually *be* a greeting card store around anymore, and then I remembered that I could buy a card at the supermarket.

Historians suggest that it may have been the early Chinese who started the trend with good will messages for the New Year. Egyptians sent greetings on ancient scrolls. In Europe, Christmas cards were the first greetings sent by mail. It wasn't until 1870 that a German immigrant in the United States began printing greeting cards and a new industry was born.

So, what *is* the state of the greeting card industry today? Would it surprise you that a report released by Global Industry Analysts, Inc. predicts that it will soon be a \$30.4 billion industry? I certainly wonder how that could be possible when I haven't gotten a greeting card in a very long time, not even one for my birthday.

These days, there are many alternatives to greeting cards such as e-cards and do-it-yourself cards. On top of that, when the economy is bad, not buying and mailing cards is an easy way to save money, and environmentalists are always reminding us to save paper. With all of these obstacles, you certainly wouldn't think the industry would be growing. So, what's the answer to these challenges? Diversify and print cards on recycled paper. We can now buy cards for cultural holidays like

Kwanzaa in Africa and Day of the Dead in Mexico. Koreans buy Valentine's cards and Canadians buy cards for Chinese New Year. We can send a card when someone retires, gets divorced, gets accepted to college, or adopts a baby. There are no constraints on the possibilities for new greeting cards.

Was I surprised by my findings? Yes. The Asia-Pacific region is the fastest growing market and the UK is the largest, most developed greeting card market in the world. Sending a card by mail may not seem very efficient in today's high-tech world, but people really get a good feeling from sending and receiving cards. I had to send a sympathy card when a friend's father died. No other form of communication seemed to have the right personal tone. I'll admit, while I was looking for the get-well card for my boss, I picked up a box of Christmas cards too. Remember how opening cards added to the holiday spirit? I'd like to bring that smile to my friends' faces this year.

Source: taken from the exam materials of Q Skills for Success Listening and Speaking

Appendix M. The Audio script of Part-2 (Content Audio) in Listening Test-2

Risk Takers

Do you ever wonder why people take risks to perform dangerous tricks?

What's the point of jumping over 20 parked cars on a motorcycle? It seems that many risk-takers do it for the fame. They live for the excitement, but they want to make sure they have an audience. Robert Thomas Jr., a reporter for *The New York Times*, wrote about "Lawbreakers We Have Known and Loved." He describes those whose daring and sometimes illegal acts capture our imagination. Why? According to Thomas, it is because they try to prove nothing is impossible; they are not held back by authority. They show a spirit of independence.

Thomas investigated why Keron, a 16-year-old immigrant from Trinidad, entered a New York City subway train and drove off with it. He begins with some background on other risk-takers who performed in front of crowds of anxious New Yorkers. One woman walked 1,600 feet on an unfinished bridge above the windy East River in 1901. To add to the danger, she was dressed in a long skirt. A French high-wire artist walked on a thin cable 1,350 feet above the ground from one World Trade Center tower to the other in 1974. Three years later, to beat a previous record, a mountain climber went up the side of all 110 stories of the South Tower. These risk-takers are usually not punished. The mountain climber, for example, had to pay a fine of just \$1.10, one cent for each story of the South Tower.

Thomas interviewed a psychologist, Dr. Joyce Brothers, to find out why these foolish acts fascinate us. Dr. Brothers was famous for her advice shows on American television and radio. She explained that we admire these crazy performers because "Everybody has fantasies. They actually go out and do them." She pointed out that

these acts have to go against the rules and be unscheduled and unexpected in order to really catch the public interest. Brothers added that we root for these people because our own lives are filled with too many rules.

Getting back to the story of the young man who stole the train, Thomas thinks Keron may have gone too far. It is one thing, he writes, to put one's own life at risk for a challenge. It is another to threaten the lives of others. Does it make a difference that Keron proved to be a good train operator? He drove 45 miles and made 85 careful stops on a journey that lasted three and a half hours. He was caught only when he set off a signal for going around a curve too fast. Some believe he should be arrested. Others think he deserves a chance at his real dream. He wants to attend school to learn to operate trains professionally. Dr. Brothers said she would be happy to ride a train driven by Keron. Would you?

Source: taken from the exam materials of Q Skills for Success Listening and Speaking

Appendix N. The Audio script of Part-3 (Context Video) in Listening Test-2

Climate Change

Hello, my name is Ian Burton. I am a professor at the University of Toronto in Canada where I study climate change, in the effects of climate change on the environment, on the people, and on the animals. First, I want to thank you for your interest in Hauser bears and the campaign to save and improve the bears all over the world.

Since I am from Canada, let me first talk about polar bears because in the Canadian Arctic, the ice is melting, and the sea level is rising and the habitat for the bears is shrinking. The places where they go out on the ice to catch seals and fish are becoming smaller and less accessible to them and so they are getting thinner and they are having fewer baby polar bears and their population is going down, and now they have been designated as a species at risk.

Let me tell you something funny about polar bears, however, and that is some people think that there are more of them and that is because the polar bears are moving south closer to a more human beings living. To those people it appears that there are more polar bears because they see more around, coming to hunt around their houses, look for trash and rubbish that had been thrown out by people because they are hungry and they cannot find enough food in the wild. So the irony is unless you do a proper monitoring of the bears well from the air by other means it appears to some people that there are more of them when in fact there are less.

But having talked about bears let me go on talk about other species in the Amazon rainforest, in grasslands in Africa, in mountains, in coastal areas. All over the

world species are under threat because their environment, their climate is changing rapidly and changing rapidly, of course, because of human activities especially putting carbon dioxide and other greenhouse gasses into the atmosphere which making the world warmer and is also destabilizing the climate so that not only do we have a warmer climate generally speaking but we have more extremes of climate. So, there will be more floods and more tropical storms and cyclones as well as sea level rise and droughts.

But all these are having big effects on animals, on plants and on the biosphere which are essentially the basis of human life on Earth because human life cannot exist without the biosphere to support it. If humans were disappeared, some remains of the biosphere would still be there. So it is very important that we protect our environment and biodiversity and what better place to start with bears in particular with polar bears.

Thank you.

Source: available at <https://www.youtube.com/watch?v=YXdAQmx8zB8>

Appendix O. The Audio script of Part-4 (Content Video) in Listening Test-2

Earthquake

April 18, 1906 California San Andreas Fault snaps, shaking San Francisco for nearly sixty terrifying seconds. When the trembling stops, the disaster is only beginning. Gas lines ruptured, setting off massive fires, some 700 people died, most of the city was reduced to ruins. This trembling of the ground caused when masses of rock suddenly shift below the earth's surface is called an earthquake.

Hundreds of little earthquakes shake the planet every day but most pass unnoticed. They usually occur along the boundaries of the thin plates that cover the Earth like an eggshell. Driven by the heat deep within the Earth's core, the plates grind against each other along lines called faults. When the plate's motion is blocked, stress builds up. Finally, the fault gives way. The released energy races through the Earth in the form of seismic waves.

Scientists record these waves on a device called a seismograph. These zigzag lines show the strength of various seismic waves. Using the line, scientists grade earthquakes on the Richter scale. For a quick to measure one number higher on a Richter scale, it must release about 30 times as much energy as the number below it. Every year about a hundred thousand earthquakes rumble through the ground hard enough for people to feel them. Of these only about a thousand are strong enough to damage property. But a powerful earthquake can be devastating. On average about 10,000 people die each year as a result of earthquakes.

The greatest recorded earthquake ever to hit North America measured a massive nine point two. The Tremble struck Alaska on March 28, 1964. A camera on

board a ship docked in Valdez recorded the draining of the entire harbor as a chasm opened up on the sea floor. There is no stopping the surface of the Earth from changing and moving. So engineers are focusing on ways to create better buildings, highways and bridges. Structures that will remain safe and stay in one piece. Next time the Earth begins to shape.

Source: available at <https://www.youtube.com/watch?v=VSgB1IW6O4>



Appendix P. The Audio script of Part-5 (Context Audio) in Listening Test-2

Driving Age

In most countries, getting a driver's license is still a symbol of the transition from childhood into the adult world. Teenagers can be in charge and don't need to be driven around by their parents. I concentrated on the differences in requirements for this important milestone.

The driving age varies from country to country, and within large countries like the United States, from state to state, although 18 is the most common age for legal driving around the world. I was surprised to find that the country with the youngest age is Ethiopia, where 14-year-olds can get a full license. New Zealand has the next youngest, which is 15. In two countries, the United States and Canada, the legal age varies by state. In Alberta, Canada, a 14-year-old can get a learner's permit and drive with an adult. Only an 18-year-old can have a full license there, though. In the US, the state of South Dakota allows teenagers to have an unrestricted license at the age of 14 years, 3 months. In most other states, the age is 16, with some restrictions. In California, for example, new drivers cannot drive friends around for the first 6 months. And if you are under 18, in New York City, you cannot drive with a junior license at all! My advice is to check with the local authorities, since driving regulations change frequently.

In my interviews, I was surprised to find out that in Japan, a driver's license costs over \$4000! Hungary seems to have the most barriers to teenage driving. Future drivers have to pass a general physical exam and a first aid test. Then, the practical driving test covers everything from mechanics to road signs. Teenagers also have to take over 20 hours of lessons with a driving instructor before their road test.

Here in the US, a license is generally much easier to get with low fees and few requirements.

Many young teenagers I interviewed couldn't wait to get their licenses. They said they'd be happy to run errands for their parents or drive their siblings around. For those over the age of 18, though, the excitement had quickly worn off. And, insurance companies and sociologists are noticing a surprising trend. It seems that even those who are eligible to get a license often put off what used to be such an important marker of adulthood. Although California is a car-driven society, I have found a number of reasons for this delay. First, cars and insurance are very expensive. Second, parents today seem overprotective of teenagers and don't see driving them around as a burden. Strangely, the teenagers don't seem to mind either. Third, some teenagers just like the carefree feeling of not having a car and others don't want the pressure of having to drive friends around.

Things have certainly changed since I was 16!

Source: taken from the exam materials of Q Skills for Success Listening and Speaking

LISTENING TEST-1

CIRCLE YOUR LEVEL

A2

B1

B2

NAME/SURNAME:.....

CLASS:.....

A. PART-1:/ 10

B. PART-2:/ 12

C. PART-3:/ 14

D. PART-4:/ 13

E. PART-5:/ 10

A. PART-1: You are going to listen to a lecture about “People who made a great impact on our society”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. Who receives awards from different organizations every year?

- a) only famous people
- b) only ordinary people
- c) both famous and ordinary people

2. Who is Derreck Kayongo?

- a) a refugee from Uganda
- b) a hotel worker in Africa
- c) a worker for CARE in Haiti

3. What does the Global Soap Project do?

- a) It sells soap made by African children to hotels
- b) It takes soap from hotels and distributes it to children
- c) It collects money for recycled soap and uses it for medicine

4. As of 2012, how many soap bars has GSP given away?

- a) 215,000
- b) 225,000
- c) 250,000

5. Which country is **NOT** included in this project?

- a) Syria
- b) Haiti
- c) Honduras

6. Oil spilling caused _____.

- a) a crash of oil tankers
- b) a change in Francis’ life
- c) death of birds and sea life

7. John Francis felt himself responsible for oil spilling because he _____.
- a) used gas vehicles too
 - b) owned an oil company
 - c) sold gas to other countries
8. In 1970, John Francis decided to _____.
- a) work for animal rights
 - b) stop using his vehicles
 - c) improve human health
9. John Francis walked across the United States on behalf of _____.
- a) the environment
 - b) global poverty
 - c) the killed animals
10. What is the main idea of this lecture?
- a) Ordinary people can make extraordinary work.
 - b) People can build organizations in different countries.
 - c) Famous people can make an impact on organization

B. PART-2: You are going to listen to a talk about “McDonald’s”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–12.

1. How many restaurants are there according to the company’s official website?
 - a) 13,400
 - b) 30,400
 - c) 34,000

2. How many customers visit McDonald’s every day?
 - a) 7 million
 - b) 17 million
 - c) 70 million

3. When McDonald’s first opened its restaurant, it aimed for _____.
 - a) teenagers
 - b) children
 - c) adults

4. Nutritionists blame the company for _____.
 - a) high obesity rates
 - b) low quality of food
 - c) not understanding customers’ needs

5. Customers follow the main trend by choosing _____.
 - a) taste
 - b) nutrition
 - c) both taste and nutrition

6. How do fast food chains attract customers?
 - a) by offering lower food prices
 - b) by choosing taste over quality
 - c) by selling food in high quantities

7. New York was the first city to _____.
- a) open the first McDonald's restaurant
 - b) put calorie counts on McDonald's menus
 - c) provide healthy food in McDonald's restaurants
8. Presenting calorie counts on the menus showed that customers_____.
- a) try to follow healthy guidelines
 - b) don't care about them in fast food chains
 - c) are confused about which food to eat
9. According to a survey at NYU, 54% of the people _____.
- a) eat fewer calories everyday
 - b) notice the calories on the menus
 - c) think that fast food is unhealthy
10. How much advertising time does McDonald's spend on promoting salads?
- a) two thirds
 - b) one sevenths
 - c) one sixths
11. The salad example proves that customers_____.
- a) don't get enough information about salads
 - b) prefer healthy food on every occasion
 - c) choose to eat unhealthy food
12. According to the speaker, fast food restaurants_____.
- a) serve their customers' needs
 - b) must be closed immediately
 - c) should change their menus

C. PART-3: You are going to watch a video lecture about “Cigarette Smoking”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–14.

1. What is one of the biggest global problems of public health?
 - a) cancer
 - b) tuberculosis
 - c) heart disease

2. Sir Walter Raleigh discovered tobacco in the _____.
 - a) African colonies
 - b) European colonies
 - c) American colonies

3. Which substance in tobacco drives cigarette smoking?
 - a) carbon
 - b) toxics
 - c) nicotine

4. Which disease caused by smoking is **NOT** mentioned in the talk?
 - a) lung cancer
 - b) diabetes
 - c) hypertension

5. The smoking first became popular at the end of the _____.
 - a) 19th century
 - b) 18th century
 - c) 17th century

6. Smoking was first popular among _____.
 - a) the rich
 - b) rich men
 - c) only men

7. Smoking reached its height in the western world in the _____.
 - a) 1930s
 - b) 1940s
 - c) 1950s

8. What is the percentage of smokers in China?
- a) 20%
 - b) 52%
 - c) 75%
9. What is the difference between the smokers in the UK or the U.S. and China?
- a) China has more smokers than UK and the US have
 - b) UK and the US have more smokers than China has
 - c) UK, the US and China have equal numbers of smokers
10. The first scientific evidence on cigarette smoking was published in _____.
- a) 1942
 - b) 1952
 - c) 1962
11. Where is Dolan Hill from?
- a) Britain
 - b) the U.S.
 - c) China
12. According to the study done by Dolan Hill, there is a link between _____.
- a) poor health and smoking
 - b) lung cancer and smoking
 - c) poverty and smoking
13. After the war against smoking in some countries, its rates are _____.
- a) increasing rapidly
 - b) remaining the same
 - c) falling gradually
14. What is the main idea of the talk?
- a) Cigarette smoking is still a major problem in the world.
 - b) Cigarette smoking affects economy in the world.
 - c) A lot of research on cigarette smoking have been done in the world.

D. PART-4: You are going to watch a video lecture about “Global Warming”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–13.

1. How much has the Earth’s temperature risen in the world?
 - a) from about 0.2 to 0.4 F⁰
 - b) from about 1.2 to 1.4 F⁰
 - c) from about 2.1 to 4.1 F⁰

2. According to the scientists, what is the main cause of this temperature rise?
 - a) growing population
 - b) human activity
 - c) position of the sun

3. Burning fossil fuels began after the _____.
 - a) industrial revolution
 - b) economic revolution
 - c) environmental revolution

4. _____ is **NOT** given as an example of fossil fuels.
 - a) Natural gas
 - b) Coal
 - c) Oil

5. What do greenhouse gases do?
 - a) They keep the heat near the earth.
 - b) They reflect the heat back to the space.
 - c) They clean the atmosphere.

6. The year 1998 was recorded as _____ in the history of the Earth.
 - a) the coolest year
 - b) the most polluted year
 - c) the warmest year

7. Carbon dioxide and methane gases in the atmosphere have hit their highest levels in the past _____.
- a) 420, 000 years
 - b) 404, 000 years
 - c) 402, 000 years
8. How much has the Arctic Sea ice decreased according to NASA studies?
- a) by 30%
 - b) by 20%
 - c) by 10%
9. The concentration of greenhouse gases in the atmosphere rises because of _____.
- a) using too much water
 - b) using a lot of lands for farming
 - c) using fossil fuels
10. How much will temperature increase by the end of the century?
- a) by 2 to 10 F⁰
 - b) by 10 to 12 F⁰
 - c) by 12 to 20 F⁰
11. Which environmental change of the Earth is **NOT** mentioned?
- a) Drought
 - b) Flood
 - c) Earthquake
12. Which way of reducing the impact of global warming is **NOT** mentioned?
- a) switching the light bulbs
 - b) using less air conditioning
 - c) driving fewer miles
13. These simple changes may help keep the Earth _____.
- a) cleaner
 - b) safer
 - c) cooler

E. PART-5: You are going to listen to a lecture on “First Impressions of Food”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. What can give a positive feeling about a homemade breakfast?
 - a) a smell of the coffee
 - b) a taste of instant coffee
 - c) a sight of a cup of coffee

2. The example of ‘the morning coffee’ proves that _____.
 - a) we usually drink coffee in the morning
 - b) coffee smell may influence our first impression
 - c) coffee tastes as good as it smells

3. What kind of food causes negative impression?
 - a) ‘bigmouth sleeper’ fish
 - b) boiled vegetables
 - c) fried meals

4. Why do some people refuse to eat sushi?
 - a) They think that it isn’t delicious.
 - b) They don’t like the smell of it.
 - c) They dislike the way it feels in their mouth.

5. Covering meat with an unfamiliar sauce makes _____.
 - a) the meat more delicious
 - b) people suspicious of the meat
 - c) chefs more popular

6. Chefs try to increase the sales of their meals by _____.
 - a) serving food from all around the world.
 - b) printing photos of food on the menu.
 - c) presenting their food that looks good.

7. We all think that taste is the _____ sense in forming our first impression of food.
 - a) most important
 - b) lastly used
 - c) least important

8. Before we taste the food, _____.

- a) our impression is affected by its name
- b) other senses influence us strongly
- c) we smell it to know if it is spicy

9. Where is durian fruit grown?

- a) in Europe
- b) in Asia
- c) in Australia

10. What is the main idea of the talk?

- a) Many senses contribute to our first impressions of foods
- b) Our first impression of food is based on taste.
- c) The sight is becoming less important in our impression of food.

LISTENING TEST-2

CIRCLE YOUR LEVEL

A2

B1

B2

NAME/SURNAME:

CLASS:.....

A. PART-1 :/ 9

B. PART-2 :/ 16

C. PART-3 :/ 11

D. PART-4 :/ 13

E. PART-5 :/ 13

A. PART-1: You are going to listen to a talk about “Greeting Cards”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–9.

1. What greeting cards were sent by mail first?
 - a) Christmas
 - b) New Year
 - c) Birthday

2. Who printed the greeting cards first?
 - a) the Germans
 - b) the Americans
 - c) the Egyptians

3. What happened in the United States?
 - a) The first greeting card store was opened.
 - b) Greeting cards were sold in the supermarkets.
 - c) Printing greeting cards began on a large scale.

4. According to a report, greeting card industry costs _____.
 - a) \$ 30.4 billion
 - b) \$ 34.0 billion
 - c) \$ 34.4 billion

5. What surprises the speaker?
 - a) He still receives so many greeting cards.
 - b) The greeting card industry is still growing.
 - c) There are no more greeting card stores.

6. Which special occasion is **NOT** mentioned in the talk?
 - a) getting retired
 - b) getting divorced
 - c) getting married

7. According to the speaker, which statement is true today?
- a) The Asia-Pacific region is the fastest growing greeting card market.
 - b) No one really appreciates getting greeting cards anymore.
 - c) There is less variety of occasions for sending greeting cards.
8. _____ has the most developed greeting card market in the world.
- a) the United Kingdom
 - b) the Asia-Pacific region
 - c) the United States
9. What is the main idea of the talk?
- a) Which occasions are good for sending greeting cards.
 - b) Why sending greeting cards is very important.
 - c) How the greeting card industry is growing

B. PART-2: You are going to listen to a report about “Risk Takers”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–16.

1. Why do people take risks to perform dangerous tricks?
 - a) for the anxiety
 - b) for the fame
 - c) for the money

2. Robert Thomes Jr. is a _____.
 - a) writer
 - b) reporter
 - c) risk-taker

3. According to Robert Thomes Jr, the purpose of risk-takers is to _____.
 - a) prove nothing is impossible
 - b) earn a lot of money
 - c) teach how to take risks

4. What did Keron, the 16-year old boy from Trinidad, do?
 - a) Walked across an unfinished bridge.
 - b) Climbed New York's tallest building.
 - c) Drove a train without permission.

5. One woman walked _____ feet above the ground in 1901.
 - a) 1006
 - b) 1060
 - c) 1600

6. What did the woman wear in 1901?
 - a) a long dress
 - b) a long skirt
 - c) a long coat

7. A French artist walked _____ feet above the ground.
 - a) 1305
 - b) 1315
 - c) 1350

8. When did a French artist walk on a thin cable?
 - a) in 1971
 - b) in 1974
 - c) in 1977

9. Three years later, a mountain climber went up 110 stories of a tower to _____.
- a) beat the previous record
 - b) show a spirit of independence
 - c) have more audience
10. The risk-takers mentioned in Thomas's article were _____.
- a) often arrested for their actions
 - b) not usually punished for their acts
 - c) never caught while committing a crime
11. Who was Dr. Joyce Brothers?
- a) a writer
 - b) a TV reporter
 - c) a psychologist
12. People admire these crazy performers because they _____.
- a) fulfill their fantasies
 - b) are exciting to watch
 - c) don't follow the rules
13. How many miles did Keron drive the train?
- a) 35
 - b) 45
 - c) 85
14. How long did Keron's journey take by train?
- a) an hour and a half
 - b) two hours and a half
 - c) three hours and a half
15. Why did Keron steal the train?
- a) He wanted to show off in front of people.
 - b) It was his biggest childhood dream.
 - c) He wanted to prove that trains are not safe.
16. What is the main idea of the talk?
- a) Everyone wants to prove that they can take risks.
 - b) Unusual people take risks all around the world.
 - c) Some people love to take risks for different reasons.

C. PART-3: You are going to watch a video talk about “Climate Change”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–11.

1. What is the speaker`s job?
 - a) an editor
 - b) an academic
 - c) an author

2. The aim of “**Hauser Bear**” campaign is to _____.
 - a) provide food for hungry bears
 - b) create an area where bears live
 - c) increase the numbers of bears

3. Which is **NOT** mentioned as a result of the climate change in Canada?
 - a) the level of the sea is rising
 - b) the prices of foods are increasing
 - c) the living areas for bears are getting smaller

4. The population of the polar bears decreases because _____.
 - a) they can` t find food
 - b) people hunt them
 - c) the weather is cold

5. People think that there are more polar bears because _____.
 - a) bears are going to places where humans live
 - b) people are moving to lands where bears live
 - c) hunters see a lot of bears in the north

6. What is the problem that bears have in the wildlife?
 - a) They don` t have any place to live.
 - b) They cannot find enough food.
 - c) They aren` t protected by people.

7. Which areas are **NOT** mentioned as under threat?
- a) deserts
 - b) grasslands
 - c) rainforests
8. Species in the world are under threat because _____.
- a) extreme weather is destroying their habitat
 - b) the climate is changing very fast
 - c) the sea levels are rising rapidly
9. What kind of human activity is mentioned as a reason of climate change?
- a) Putting dangerous gases into the atmosphere.
 - b) Cutting a lot of trees in the big forests.
 - c) Building many houses on the farm lands.
10. Which is **NOT** mentioned as a result of extreme weather conditions?
- a) floods
 - b) drought
 - c) famine
11. What is the main idea of the talk?
- a) The main causes of global warming in the world.
 - b) The effects of the climate change on the environment.
 - c) The results of extreme weather conditions on the Earth.

D. PART-4: You are going to watch a video lecture about “Earthquake”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–13.

1. When did the earthquake happen?
 - a) April 16, 1606
 - b) April 19, 1806
 - c) April 18, 1906

2. Which was **NOT** mentioned as the effect of the earthquake?
 - a) damaged bridges
 - b) destroyed gas line
 - c) very big fires

3. How many people died because of the earthquake?
 - a) 700
 - b) 1700
 - c) 7000

4. How many little earthquakes shake the planet every day?
 - a) a hundred
 - b) hundreds
 - c) over a hundred

5. Most of the little earthquakes are _____.
 - a) happening fast
 - b) felt by a few
 - c) not noticed

6. What does a seismograph **NOT** do with earthquake waves?
 - a) It records them.
 - b) It shows their levels.
 - c) It warns us about them.

7. How many earthquakes are felt by people every year?
- a) less than 100,000
 - b) about 100,000
 - c) over 100,000
8. Only about 1000 earthquakes _____.
- a) do not cause damage
 - b) occur in mega cities
 - c) destroy people`s property
9. How many people die because of earthquakes each year?
- a) less than 10,000
 - b) about 10,000
 - c) over 10,000
10. The biggest earthquake was recorded in _____.
- a) North America
 - b) Eastern Asia
 - c) Southern Africa
11. When did the last earthquake happen in Alaska?
- a) 1674
 - b) 1794
 - c) 1964
12. Which was **NOT** mentioned as idea engineers focus on?
- a) safe structures
 - b) new cities
 - c) better bridges
13. The video includes the information about _____.
- a) the effects of earthquake on people
 - b) the reasons for earthquakes to occur
 - c) the facts about earthquakes in the world

E. PART-5: You are going to listen to a talk on “Driving Age” Listen carefully and choose the correct answer (a, b or c) for Questions 1–13.

1. Having a driver’s license is important because _____.
 - a) some countries don’t have it
 - b) it is a milestone representing adulthood
 - c) there are a lot of requirements for having it

2. What does the age of 18 represent?
 - a) The average age of having a license in the United States
 - b) The minimum age of having a license in a few countries
 - c) The common age of having a license in most countries

3. What is the age at which you can have a driving license in New Zealand?
 - a) 14
 - b) 15
 - c) 18

4. In the USA and Canada, the age of having a driving license _____.
 - a) changes in each state
 - b) will be increased by law
 - c) remains the same

5. In which state you can drive with no restriction at the age of 14?
 - a) California
 - b) South Dakota
 - c) New York

6. In California, which is mentioned as a way to restrict junior drivers?
 - a) not driving with cell phones.
 - b) not driving alone outside the city
 - c) not driving with their friends around

7. Where cannot you drive with a junior license if you are under 18?
 - a) in New York City
 - b) in California
 - c) in South Dakota

- 8.** What surprised Meghan about Japan`s driving regulations?
- a) the age of teenagers that could drive
 - b) the cost of Japanese drivers` licenses
 - c) the number of barriers to teenage driving
- 9.** In which country do teenagers have to take 20 hours of driving lessons?
- a) Hungary
 - b) Japan
 - c) the U.S.
- 10.** In which country having a driving license is the easiest?
- a) Hungary
 - b) Japan
 - c) the USA
- 11.** Teenagers can`t wait to have a license because they wanted to _____.
- a) buy their own car
 - b) drive their family around
 - c) prove that they are adults
- 12.** People over the age of 18 do not have a driving license because _____.
- a) traffic accidents are increasing
 - b) the prices of cars are very high
 - c) parents don`t support them financially
- 13.** What is the main idea of the talk?
- a) Why teenagers cannot get driver`s license in some countries.
 - b) The main reasons why teenagers get their driver`s license.
 - c) How teenagers can get driver`s license around the world.

LISTENING TEST-1

CIRCLE YOUR LEVEL

A2

B1

B2

NAME/SURNAME:

CLASS:

A. PART-1:/ 10

B. PART-2:/ 10

C. PART-3:/ 10

D. PART-4:/ 10

E. PART-5:/ 10

A. PART-1: You are going to listen to a lecture about “People who made a great impact on our society”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. Who receives awards from different organizations every year?
 - a) only famous people
 - b) only ordinary people
 - c) both famous and ordinary people

2. Who is Derreck Kayongo?
 - a) a refugee from Uganda
 - b) a hotel worker in Africa
 - c) a worker for CARE in Haiti

3. What does the Global Soap Project do?
 - a) It sells soap made by African children to hotels
 - b) It takes soap from hotels and distributes it to children
 - c) It collects money for recycled soap and uses it for medicine

4. As of 2012, how many soap bars has GSP given away?
 - a) 215,000
 - b) 225,000
 - c) 250,000

5. Which country is **NOT** included in this project?
 - a) Syria
 - b) Haiti
 - c) Honduras

6. Oil spilling caused _____.
 - a) a crash of oil tankers
 - b) a change in Francis’ life
 - c) death of birds and sea life

7. John Francis felt himself responsible for oil spilling because he _____.
 - a) used gas vehicles too
 - b) owned an oil company
 - c) sold gas to other countries

- 8.** In 1970, John Francis decided to _____.
- a) work for animal rights
 - b) stop using his vehicles
 - c) improve human health
- 9.** John Francis walked across the United States on behalf of _____.
- a) the environment
 - b) global poverty
 - c) the killed animals
- 10.** What is the main idea of this lecture?
- a) Ordinary people can make extraordinary work.
 - b) People can build organizations in different countries.
 - c) Famous people can make an impact on organizations.

B. PART-2: You are going to listen to a talk about “McDonald’s”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. How many restaurants are there according to the company’s official website?
 - a) 13,400
 - b) 30,400
 - c) 34,000

2. How many customers visit McDonald’s every day?
 - a) 7 million
 - b) 17 million
 - c) 70 million

3. When McDonald’s first opened its restaurant, it aimed for _____.
 - a) teenagers
 - b) children
 - c) adults

4. Nutritionists blame the company for _____.
 - a) high obesity rates
 - b) low quality of food
 - c) not understanding customers’ needs

5. Customers follow the main trend by choosing _____.
 - a) taste
 - b) nutrition
 - c) both taste and nutrition

6. How do fast food chains attract customers?
 - a) by offering lower food prices
 - b) by choosing taste over quality
 - c) by selling food in high quantities

7. New York was the first city to _____.
- a) open the first McDonald's restaurant
 - b) put calorie counts on McDonald's menus
 - c) provide healthy food in McDonald's restaurants
8. Presenting calorie counts on the menus showed that customers_____.
- a) try to follow healthy guidelines
 - b) don't care about them in fast food chains
 - c) are confused about which food to eat
9. According to a survey at NYU, 54% of the people _____.
- a) eat fewer calories everyday
 - b) notice the calories on the menus
 - c) think that fast food is unhealthy
10. The salad example proves that customers_____.
- a) don't get enough information about salads
 - b) prefer healthy food on every occasion
 - c) choose to eat unhealthy food

C. PART-3: You are going to watch a video lecture about “Cigarette Smoking”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. What is one of the biggest global problems of public health?
 - a) cancer
 - b) tuberculosis
 - c) heart disease

2. Which disease caused by smoking is **NOT** mentioned in the talk?
 - a) lung cancer
 - b) diabetes
 - c) hypertension

3. The smoking first became popular at the end of the _____.
 - a) 19th century
 - b) 18th century
 - c) 17th century

4. Smoking reached its height in the western world in the _____.
 - a) 1930s
 - b) 1940s
 - c) 1950s

5. What is the percentage of smokers in China?
 - a) 20%
 - b) 52%
 - c) 75%

6. What is the difference between the smokers in the UK or the U.S. and China?
 - a) China has more smokers than UK and the US have
 - b) UK and the US have more smokers than China has
 - c) UK, the US and China have equal numbers of smokers

7. The first scientific evidence on cigarette smoking was published in _____.
 - a) 1942
 - b) 1952
 - c) 1962

- 8.** Where is Dolan Hill from?
- a) Britain
 - b) the U.S.
 - c) China
- 9.** According to the study done by Dolan Hill, there is a link between _____.
- a) poor health and smoking
 - b) lung cancer and smoking
 - c) poverty and smoking
- 10.** What is the main idea of the talk?
- a) Cigarette smoking is still a major problem in the world.
 - b) Cigarette smoking affects economy in the world.
 - c) A lot of research on cigarette smoking have been done in the world.

D. PART-4: You are going to watch a video lecture about “Global Warming”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. According to the scientists, what is the main cause of this temperature rise?
 - a) growing population
 - b) human activity
 - c) position of the sun

2. Burning fossil fuels began after the _____.
 - a) industrial revolution
 - b) economic revolution
 - c) environmental revolution

3. _____ is **NOT** given as an example of fossil fuels.
 - a) Natural gas
 - b) Coal
 - c) Oil

4. The year 1998 was recorded as _____ in the history of the Earth.
 - a) the coolest year
 - b) the most polluted year
 - c) the warmest year

5. How much has the Arctic Sea ice decreased according to NASA studies?
 - a) by 30%
 - b) by 20%
 - c) by 10%

6. The concentration of greenhouse gases in the atmosphere rises because of _____.
 - a) using too much water
 - b) using a lot of lands for farming
 - c) using fossil fuels

7. How much will temperature increase by the end of the century?
- a) by 2 to 10 F⁰
 - b) by 10 to 12 F⁰
 - c) by 12 to 20 F⁰
8. Which environmental change of the Earth is **NOT** mentioned?
- a) Drought
 - b) Flood
 - c) Earthquake
9. Which way of reducing the impact of global warming is **NOT** mentioned?
- a) switching the light bulbs
 - b) using less air conditioning
 - c) driving fewer miles
10. These simple changes may help keep the Earth _____.
- a) cleaner
 - b) safer
 - c) cooler

E. PART-5: You are going to listen to a lecture on “First Impressions of Food”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–10.

1. What can give a positive feeling about a homemade breakfast?
 - a) a smell of the coffee
 - b) a taste of instant coffee
 - c) a sight of a cup of coffee

2. The example of ‘the morning coffee’ proves that _____.
 - a) we usually drink coffee in the morning
 - b) coffee smell may influence our first impression
 - c) coffee tastes as good as it smells

3. What kind of food causes negative impression?
 - a) ‘bigmouth sleeper’ fish
 - b) boiled vegetables
 - c) fried meals

4. Why do some people refuse to eat sushi?
 - a) They think that it isn’t delicious.
 - b) They don’t like the smell of it.
 - c) They dislike the way it feels in their mouth.

5. Covering meat with an unfamiliar sauce makes _____.
 - a) the meat more delicious
 - b) people suspicious of the meat
 - c) chefs more popular

6. Chefs try to increase the sales of their meals by _____.
 - a) serving food from all around the world.
 - b) printing photos of food on the menu.
 - c) presenting their food that looks good.

7. We all think that taste is the _____ sense in forming our first impression of food.
 - a) most important
 - b) lastly used
 - c) least important

8. Before we taste the food, _____.

- a) our impression is affected by its name
- b) other senses influence us strongly
- c) we smell it to know if it is spicy

9. Where is durian fruit grown?

- a) in Europe
- b) in Asia
- c) in Australia

10. What is the main idea of the talk?

- a) Many senses contribute to our first impressions of foods
- b) Our first impression of food is based on taste.
- c) The sight is becoming less important in our impression of food.

LISTENING TEST-2

CIRCLE YOUR LEVEL

A2

B1

B2

NAME/SURNAME:

CLASS:.....

A. PART-1 :/ 9

B. PART-2 :/ 9

C. PART-3 :/ 9

D. PART-4 :/ 9

E. PART-5 :/ 9

A. PART-1: You are going to listen to a talk about “Greeting Cards”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–9.

1. What greeting cards were sent by mail first?
 - a) Christmas
 - b) New Year
 - c) Birthday

2. Who printed the greeting cards first?
 - a) the Germans
 - b) the Americans
 - c) the Egyptians

3. What happened in the United States?
 - a) The first greeting card store was opened.
 - b) Greeting cards were sold in the supermarkets.
 - c) Printing greeting cards began on a large scale.

4. According to a report, greeting card industry costs _____.
 - a) \$ 30.4 billion
 - b) \$ 34.0 billion
 - c) \$ 34.4 billion

5. What surprises the speaker?
 - a) He still receives so many greeting cards.
 - b) The greeting card industry is still growing.
 - c) There are no more greeting card stores.

6. Which special occasion is **NOT** mentioned in the talk?
 - a) getting retired
 - b) getting divorced
 - c) getting married

7. According to the speaker, which statement is true today?
- a) The Asia-Pacific region is the fastest growing greeting card market.
 - b) No one really appreciates getting greeting cards anymore.
 - c) There is less variety of occasions for sending greeting cards.
8. _____ has the most developed greeting card market in the world.
- a) the United Kingdom
 - b) the Asia-Pacific region
 - c) the United States
9. What is the main idea of the talk?
- a) Which occasions are good for sending greeting cards.
 - b) Why sending greeting cards is very important.
 - c) How the greeting card industry is growing

B. PART-2: You are going to listen to a report about “Risk Takers”. Listen carefully and choose the correct answer (a, b or c) for Questions 1–9.

1. Robert Thomes Jr. is a _____.
 - a) writer
 - b) reporter
 - c) risk-taker

2. What did Keron, the 16-year old boy from Trinidad, do?
 - a) Walked across an unfinished bridge.
 - b) Climbed New York's tallest building.
 - c) Drove a train without permission.

3. One woman walked _____ feet above the ground in 1901.
 - a) 1006
 - b) 1060
 - c) 1600

4. Three years later, a mountain climber went up 110 stories of a tower to _____.
 - a) beat the previous record
 - b) show a spirit of independence
 - c) have more audience

5. The risk-takers mentioned in Thomas's article were _____.
 - a) often arrested for their actions
 - b) not usually punished for their acts
 - c) never caught while committing a crime

6. How many miles did Keron drive the train?
 - a) 35
 - b) 45
 - c) 85

7. How long did Keron’s journey take by train?
 - a) an hour and a half
 - b) two hours and a half
 - c) three hours and a half

8. Why did Keron steal the train?
 - a) He wanted to show off in front of people.
 - b) It was his biggest childhood dream.
 - c) He wanted to prove that trains are not safe.

9. What is the main idea of the talk?

- a)** Everyone wants to prove that they can take risks.
- b)** Unusual people take risks all around the world.
- c)** Some people love to take risks for different reasons.



C. PART-3: You are going to watch a video talk about “Climate Change”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–9.

1. Which is **NOT** mentioned as a result of the climate change in Canada?
 - a) the level of the sea is rising
 - b) the prices of foods are increasing
 - c) the living areas for bears are getting smaller

2. The population of the polar bears decreases because _____.
 - a) they can't find food
 - b) people hunt them
 - c) the weather is cold

3. People think that there are more polar bears because _____.
 - a) bears are going to places where humans live
 - b) people are moving to lands where bears live
 - c) hunters see a lot of bears in the north

4. What is the problem that bears have in the wildlife?
 - a) They don't have any place to live.
 - b) They cannot find enough food.
 - c) They aren't protected by people.

5. Which areas are **NOT** mentioned as under threat?
 - a) deserts
 - b) grasslands
 - c) rainforests

6. Species in the world are under threat because _____.
 - a) extreme weather is destroying their habitat
 - b) the climate is changing very fast
 - c) the sea levels are rising rapidly

7. What kind of human activity is mentioned as a reason of climate change?
- a) Putting dangerous gases into the atmosphere.
 - b) Cutting a lot of trees in the big forests.
 - c) Building many houses on the farm lands.
8. Which is **NOT** mentioned as a result of extreme weather conditions?
- a) floods
 - b) drought
 - c) famine
9. What is the main idea of the talk?
- a) The main causes of global warming in the world.
 - b) The effects of the climate change on the environment.
 - c) The results of extreme weather conditions on the Earth.

D. PART-4: You are going to watch a video lecture about “Earthquake”. Watch the video carefully and choose the correct answer (a, b or c) for Questions 1–9.

1. Which was **NOT** mentioned as the effect of the earthquake?
 - a) damaged bridges
 - b) destroyed gas line
 - c) very big fires

2. How many people died because of the earthquake?
 - a) 700
 - b) 1700
 - c) 7000

3. How many little earthquakes shake the planet every day?
 - a) a hundred
 - b) hundreds
 - c) over a hundred

4. Most of the little earthquakes are _____.
 - a) happening fast
 - b) felt by a few
 - c) not noticed

5. What does a seismograph **NOT** do with earthquake waves?
 - a) It records them.
 - b) It shows their levels.
 - c) It warns us about them.

6. Only about 1000 earthquakes _____.
 - a) do not cause damage
 - b) occur in mega cities
 - c) destroy people`s property

7. How many people die because of earthquakes each year?
- a) less than 10,000
 - b) about 10,000
 - c) over 10,000
8. Which was **NOT** mentioned as idea engineers focus on?
- a) safe structures
 - b) new cities
 - c) better bridges
9. The video includes the information about _____.
- a) the effects of earthquake on people
 - b) the reasons for earthquakes to occur
 - c) the facts about earthquakes in the world

E. PART-5: You are going to listen to a talk on “Driving Age” Listen carefully and choose the correct answer (a, b or c) for Questions 1–9.

1. What does the age of 18 represent?
 - a) The average age of having a license in the United States
 - b) The minimum age of having a license in a few countries
 - c) The common age of having a license in most countries

2. In the USA and Canada, the age of having a driving license _____.
 - a) changes in each state
 - b) will be increased by law
 - c) remains the same

3. In which state you can drive with no restriction at the age of 14?
 - a) California
 - b) South Dakota
 - c) New York

4. In California, which is mentioned as a way to restrict junior drivers?
 - a) not driving with cell phones.
 - b) not driving alone outside the city
 - c) not driving with their friends around

5. Where cannot you drive with a junior license if you are under 18?
 - a) in New York City
 - b) in California
 - c) in South Dakota

6. In which country do teenagers have to take 20 hours of driving lessons?
 - a) Hungary
 - b) Japan
 - c) the U.S.

7. In which country having a driving license is the easiest?

- a) Hungary
- b) Japan
- c) the USA

8. Teenagers can't wait to have a license because they wanted to _____.

- a) buy their own car
- b) drive their family around
- c) prove that they are adults

9. What is the main idea of the talk?

- a) Why teenagers cannot get driver's license in some countries.
- b) The main reasons why teenagers get their driver's license.
- c) How teenagers can get driver's license around the world.

Appendix U. Semi-Structured Group Interview Questions

Interview Questions

1. What do you think about the listening tests such as the length of it, playing the recordings twice, the topics of the texts, the accent, and the speech rate of the recordings?
2. What do you think about the visuals in the tests?
3. Do you think visuals, which are about the topics of the recordings, are related to the topics of the texts?
4. Did visuals help you understand the texts better or less?
5. How did the presence and the absence of the visuals affect you and your understanding?
6. Which kind of visuals helped you more or less in the listening tests, context or content?
7. What kind of visuals do you prefer in listening tests, content visuals or context visuals?
8. Would you like to listen audio recordings with visuals or without visuals?
9. Which mode would you prefer in listening tests video, audio with visuals, or audio-only?
10. What do you think about the quality of the recordings?
11. Did something in audios or videos prevent you from understanding?
12. What do you think about the topics of the recordings?
13. Was there anything that you were not familiar in the recordings in terms of the topics and contents of them?
14. What do you think about the level of the questions, were they easy/normal/difficult?