

**Marmara University  
Institute of Educational Sciences  
Department of Foreign Language Education  
Division of English Language Teaching**



**THE EFFECTS OF TASK INVOLVEMENT LOAD AND INPUT TYPE  
ON FOREIGN LANGUAGE VOCABULARY LEARNING**

**Cansu KIVRAK  
(M.A. Thesis)**

**İstanbul – 2019**

**Marmara University  
Institute of Educational Sciences  
Department of Foreign Language Education  
Division of English Language Teaching**



**THE EFFECTS OF TASK INVOLVEMENT LOAD AND INPUT TYPE  
ON FOREIGN LANGUAGE VOCABULARY LEARNING**

**Cansu KIVRAK  
(M.A. Thesis)**

**SUPERVISOR  
Assist. Prof. Dilek UYGUN GÖKMEN**

**İstanbul – 2019**

**T.C.  
Marmara Üniversitesi  
Eğitim Bilimleri Enstitüsü  
Yabancı Diller Eğitimi Ana Bilim Dalı  
İngiliz Dili Eğitimi Bilim Dalı**

**GÖREV KATILIM YÜKÜ VE VERİ BİÇİMİNİN YABANCI DİLDE  
SÖZCÜK ÖĞRENİMİNE ETKİSİ**

**Cansu KIVRAK  
(Yüksek Lisans Tezi)**

**İstanbul – 2019**

**T.C.  
Marmara Üniversitesi  
Eğitim Bilimleri Enstitüsü  
Yabancı Diller Eğitimi Ana Bilim Dalı  
İngiliz Dili Eğitimi Bilim Dalı**

**GÖREV KATILIM YÜKÜ VE VERİ BİÇİMİNİN YABANCI DİLDE  
SÖZCÜK ÖĞRENİMİNE ETKİSİ**

**Cansu KIVRAK  
(Yüksek Lisans Tezi)**

**DANIŞMAN  
Dr. Öğr. Üyesi Dilek UYGUN GÖKMEN**

**İstanbul – 2019**



**All rights belong to**

**M.U. Institute of Educational Sciences.**

**© 2019**



**Tüm kullanım hakları**

**M.Ü. Eğitim Bilimleri Enstitüsü'ne aittir.**

**© 2019**

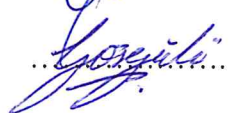
## THESIS APPROVAL

The thesis entitled “The Effects of Task Involvement Load and Input Type on Foreign Language Vocabulary Learning” completed by Cansu KIVRAK was defended successfully on 19.06.2019 and certified by examining committee members as thesis for the degree of Master of Arts.

|                   | Name Surname                      | Signature                                                                             |
|-------------------|-----------------------------------|---------------------------------------------------------------------------------------|
| THESIS SUPERVISOR | Dr. Öğr. Üyesi Dilek UYGUN GÖKMEN |    |
| JURY MEMBER       | Dr. Öğr. Üyesi Özlem KAŞLIOĞLU    |   |
| JURY MEMBER       | Dr. Öğr. Üyesi Emrah GÖRGÜLÜ      |  |

## ONAY

Cansu KIVRAK tarafından hazırlanan “Görev Katılım Yüğü ve Veri Biçiminin Yabancı Dilde Sözcük Öğrenimine Etkisi” konulu bu çalışma 19.06.2019 tarihinde yapılan savunma sınavı sonucunda jüri tarafından başarılı bulunmuş ve yüksek lisans tezi olarak kabul edilmiştir.

|               | Adı Soyadı                        | İmza                                                                                  |
|---------------|-----------------------------------|---------------------------------------------------------------------------------------|
| TEZ DANIŞMANI | Dr. Öğr. Üyesi Dilek UYGUN GÖKMEN |    |
| JÜRİ ÜYESİ    | Dr. Öğr. Üyesi Özlem KAŞLIOĞLU    |   |
| JÜRİ ÜYESİ    | Dr. Öğr. Üyesi Emrah GÖRGÜLÜ      |  |

## CURRICULUM VITAE

**NAME:** Cansu KIVRAK

**PLACE OF BIRTH:** Yıldırım, Bursa, Turkey

**DATE OF BIRTH:** 11 January 1989

### EDUCATION:

|                    |                                                                           |
|--------------------|---------------------------------------------------------------------------|
| <b>2015 – 2019</b> | M.A., English Language Teaching<br>Marmara University, İstanbul, Turkey   |
| <b>2007 – 2012</b> | B.A., Foreign Language Education<br>Boğaziçi University, İstanbul, Turkey |
| <b>2003 – 2007</b> | İbrahim Önal Science High School<br>Mustafakemalpaşa, Bursa, Turkey       |

### WORK EXPERIENCE:

|                            |                                                                         |
|----------------------------|-------------------------------------------------------------------------|
| <b>Feb 2014 – Present</b>  | Yalova University, Yalova, Turkey<br>Lecturer & Testing officer         |
| <b>Sep 2012 – Feb 2014</b> | Dumlupınar University, Kütahya, Turkey<br>Assistant director & Lecturer |

### CONFERENCE PRESENTATIONS:

Kıvrak, C. & Uygun Gökmen, D. (2019). *The effects of task-induced involvement load, input modality and language proficiency on incidental vocabulary acquisition*. GlobELT 2019 Conference. Northern Cyprus, Turkey.

Kıvrak, C. (2016). *The effects of different tasks with the same involvement load on L2 word retention*. 2<sup>nd</sup> International IATEFL Research SIG Conference: Teachers Research! Istanbul 2016. Bahçeşehir University, İstanbul, Turkey.

### CONTACT INFORMATION:

**Institution:** Yalova University, School of Foreign Languages

**E-mail:** [cansukivrk@gmail.com](mailto:cansukivrk@gmail.com)

## ÖZGEÇMİŞ

**AD-SOYAD:** Cansu KIVRAK

**DOĞUM YERİ:** Yıldırım, Bursa

**DOĞUM TARİHİ:** 11 Ocak 1989

### EĞİTİM:

|                    |                                                                       |
|--------------------|-----------------------------------------------------------------------|
| <b>2015 – 2019</b> | Marmara Üniversitesi, İstanbul<br>Yüksek Lisans, İngiliz Dili Eğitimi |
| <b>2007 – 2012</b> | Boğaziçi Üniversitesi, İstanbul<br>Lisans, İngilizce Öğretmenliği     |
| <b>2003 – 2007</b> | İbrahim Önal Fen Lisesi, Mustafakemalpaşa, Bursa                      |

### İŞ DENEYİMİ:

|                                |                                                                                                          |
|--------------------------------|----------------------------------------------------------------------------------------------------------|
| <b>Şubat 2014 – Halen</b>      | Yabancı Diller Yüksekokulu, Yalova Üniversitesi<br>Öğretim Görevlisi & Ölçme Değerlendirme Ofisi Elemanı |
| <b>Eylül 2012 – Şubat 2014</b> | Yabancı Diller Yüksekokulu, Dumlupınar Üniversitesi<br>Bölüm Başkan Yardımcısı & Öğretim Görevlisi       |

### SUNULAN BİLDİRİLER:

Kıvrak, C. & Uygun Gökmen, D. (2019). *The effects of task-induced involvement load, input modality and language proficiency on incidental vocabulary acquisition*. GlobELT 2019 Konferansı. Kuzey Kıbrıs, Türkiye.

Kıvrak, C. (2016). *The effects of different tasks with the same involvement load on L2 word retention*. 2. Uluslararası IATEFL Research SIG Konferansı: Teachers Research! İstanbul 2016. Bahçeşehir Üniversitesi, İstanbul, Türkiye.

### İLETİŞİM BİLGİLERİ:

**Görev Yaptığı Kurum:** Yabancı Diller Yüksekokulu, Yalova Üniversitesi

**E-posta:** [cansukivrk@gmail.com](mailto:cansukivrk@gmail.com)

## ABSTRACT

In this study, the purpose was to explore the effects of task-induced involvement load, input type and language proficiency on Turkish EFL learners' productive and receptive word knowledge. In other words, the present study investigated how variously loaded tasks or tasks with the same involvement load but different input modalities (written vs. audiovisual) affect incidental vocabulary acquisition at different proficiency levels. The secondary aim was to provide empirical evidence to the Involvement Load Hypothesis (Laufer & Hulstijn, 2001) by testing its assumptions regarding task efficacy and L2 vocabulary learning. For these purposes, the following research questions were addressed:

- 1) Do vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at lower-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 2) Do vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at lower-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 3) Do vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at upper-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 4) Do vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at upper-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?

The participants included 236 Turkish (lower- or upper-intermediate) EFL learners enrolled in an Intensive English Program at a preparatory school of a state university in

Turkey. Data were collected with the use of several instruments administered to the two proficiency groups in the spring term of 2016-2017 academic year: 2 video podcasts, 2 reading passages, 12 vocabulary tasks (6 per level), 2 vocabulary pre-tests and 4 productive and receptive vocabulary post-tests. The quantitative data that came from the immediate and delayed vocabulary post-tests were analyzed using SPSS 23.0.

The results showed that for productive word knowledge, the pairing of the gap filling task with the written input and the pairing of the sentence writing task with the audiovisual input were more effective for both proficiency levels. As for receptive word knowledge, the sentence writing task (higher involvement load) was more effective than the gap filling task (lower involvement load) for both proficiency levels regardless of input type, but the audiovisual input fostered more receptive word knowledge among the upper-intermediate learners. While these findings partially confirmed the predictions of the ILH for receptive word knowledge, they also highlighted the role that input modality plays in increasing productive word knowledge. Finally, the findings of the study were discussed in relation to the relevant literature and several pedagogical suggestions were made for language teachers and material developers.

**Keywords:** incidental vocabulary learning, task-induced involvement load, input type, task efficacy, vocabulary gain and retention

## ÖZET

Bu araştırmada, görev kaynaklı katılım yükü, veri biçimi ve yabancı dil seviyesinin İngilizce'yi yabancı dil olarak öğrenen Türk öğrencilerin algısal ve üretimsel sözcük bilgisi üzerindeki etkilerini araştırmak amaçlanmıştır. Başka bir deyişle, bu araştırma farklı görev katılım yüklerine sahip ya da aynı görev katılım yükü fakat farklı veri biçimlerini (yazılı ya da görsel-işitsel) içeren aktivitelerin farklı dil seviyelerinde rastlantısal sözcük edinimini nasıl etkilediğini araştırmıştır. Çalışmada ikincil olarak ise Görev Kaynaklı Katılım Yüğü Hipotezi'nin (Laufer ve Hulstijn, 2001) aktivite verimliliği ve ikinci yabancı dilde kelime öğrenimi ile ilgili varsayımları test edilerek hipoteze ampirik bulgularla katkıda bulunmak amaçlanmıştır. Bu amaçlar için, aşağıdaki araştırma soruları ele alınmıştır:

- 1) Alt-orta dil seviyesinde, daha yüksek görev katılım yüküne sahip kelime aktiviteleri daha düşük görev katılım yüküne sahip olanlara göre
  - a) daha fazla üretimsel sözcük bilgisi edinimi sağlar mı?
  - b) daha fazla algısal sözcük bilgisi edinimi sağlar mı?
- 2) Alt-orta dil seviyesinde, farklı veri biçimlerine (yazılı ya da görsel-işitsel) sahip kelime aktiviteleri
  - a) üretimsel sözcük bilgisi edinimi açısından farklılık yaratır mı?
  - b) algısal sözcük bilgisi edinimi açısından farklılık yaratır mı?
- 3) Üst-orta dil seviyesinde, daha yüksek görev katılım yüküne sahip kelime aktiviteleri daha düşük görev katılım yüküne sahip olanlara göre
  - a) daha fazla üretimsel sözcük bilgisi edinimi sağlar mı?
  - b) daha fazla algısal sözcük bilgisi edinimi sağlar mı?
- 4) Üst-orta dil seviyesinde, farklı veri biçimlerine (yazılı ya da görsel-işitsel) sahip kelime aktiviteleri
  - a) üretimsel sözcük bilgisi edinimi açısından farklılık yaratır mı?
  - b) algısal sözcük bilgisi edinimi açısından farklılık yaratır mı?

Araştırmanın örneklemini Türkiye'de bir devlet üniversitesinin hazırlık okulunda Yoğun İngilizce Programı'na kayıtlı İngilizce'yi yabancı dil olarak (alt- veya üst-orta dil seviyesinde) öğrenen 236 Türk öğrenciden oluşmaktadır. Veriler 2016-2017 akademik

yılının bahar döneminde iki dil seviyesine de uygulanan birkaç ölçeğin kullanılması ile toplanmıştır: 2 video yayını, 2 okuma parçası, 12 kelime aktivitesi (her seviye için 6), 2 kelime ön testi ve 4 üretimsel ve algısal kelime son testi. Aktivite sonrası hemen yapılan kelime son testleri ve kalıcılık kelime son testlerinden elde edilen nicel veriler SPSS 23.0 kullanılarak analiz edilmiştir.

Araştırmadan elde edilen sonuçlar, boşluk doldurma aktivitesinin yazılı veri biçimi ile, cümle yazma aktivitesinin ise görsel-işitsel veri biçimi ile eşleşmesinin üretimsel sözcük bilgisi açısından her iki dil seviyesinde de daha etkili olduğunu göstermiştir. Algısal sözcük bilgisine gelince, her iki dil seviyesi için de veri biçimi fark etmeksizin cümle yazma aktivitesinin (daha yüksek görev katılım yüklü) boşluk doldurma aktivitesinden (daha düşük görev katılım yüklü) daha etkili olduğu, fakat görsel-işitsel veri biçiminin üst-orta dil seviyesindeki öğrenciler arasında daha fazla algısal kelime edinimini sağladığı görülmüştür. Bu bulgular Görev Kaynaklı Katılım Yüklü Hipotezi'nin algısal sözcük bilgisine dair varsayımlarını kısmen doğrularken, veri biçiminin üretimsel sözcük bilgisini artırmada oynadığı role de ayrıca dikkat çekmiştir. Son olarak, araştırmanın bulguları ilgili alanyazına değinilerek tartışılmış ve dil öğretmenleri ve materyal geliştirenler için eğitimsel önerilerde bulunulmuştur.

**Anahtar kelimeler:** rastlantısal kelime öğrenimi, görev kaynaklı katılım yükü, veri biçimi, aktivite verimliliği, kelime edinimi ve kalıcılığı

## ACKNOWLEDGMENT

This thesis becomes a reality with the support and help of many people around me. Therefore, I would like to thank all of them for their huge contribution.

Foremost, I would like to express my special gratitude to my thesis advisor Assist. Prof. Dr. Dilek UYGUN GÖKMEN for her continuous support, endless patience, positive encouragement, extensive knowledge and invaluable feedback that helped me considerably in the process of writing my thesis. She was always there to guide me whenever I had a question about my research or writing.

Besides my advisor, I would like to thank the rest of my thesis committee: Assist. Prof. Dr. Özlem KAŞLIOĞLU and Assist. Prof. Dr. Emrah GÖRGÜLÜ, not only for their constructive comments but also for their enlightening questions which enabled me to widen my perspective on my research.

I am highly indebted to my fellow colleagues and their students for their participation in my study and their contribution to my data collection. Without their help, this thesis would not have been possible.

Last but not the least, I would like to dedicate my thesis to my loving mother Yasemin CAN. I must express my very profound gratitude to her for always encouraging me along this difficult journey and wholeheartedly believing in me throughout my life. This and many other accomplishments in my life would not have been possible without her.

Cansu KIVRAK

## ÖNSÖZ

Bu tez desteęi ve yardımıyla yanımda olan kişilerin sayesinde gerçekleşebildi. Bu nedenle, büyük katkılarından dolayı onlara teşekkür etmek istiyorum.

En başta, devamlı desteęi, sonsuz sabrı, pozitif teşviki, kapsamlı bilgisi ve değerli yorumlarıyla tez yazım sürecinde bana önemli ölçüde yardım eden ve araştırma ya da yazma aşamasında ne zaman bir sorum olsa bana daima rehberlik eden tez danışmanım Dr. Öğr. Üyesi Dilek UYGUN GÖKMEN'e,

Danışmanımın yanı sıra, yapıcı yorumlarının yanında aydınlatan soruları ile araştırmama olan bakış açımı genişletmemi sağlayan tez komite üyelerim Dr. Öğr. Üyesi Özlem KAŞLIOĞLU'na ve Dr. Öğr. Üyesi Emrah GÖRGÜLÜ'ye,

Araştırmama katılıp data toplamamda bana katkıda bulunarak yardımlarıyla bu tezi mümkün kılan değerli meslektaşlarım ve öğrencilerine teşekkürü bir borç bilirim.

Son ama en önemlisi, tezimi sevgi dolu annem Yasemin CAN'a ithaf etmek istiyorum. Hayatım boyunca bütün kalbiyle bana inanarak bu uzun yolculukta beni her zaman destekledięi için ona çok derin minnet duyuyorum. Hayatımdaki bu ve bütün diğer başarılarım onsuz mümkün olamazdı.

Cansu KIVRAK

## TABLE OF CONTENTS

|                                                                       |     |
|-----------------------------------------------------------------------|-----|
| THESIS APPROVAL.....                                                  | i   |
| ONAY.....                                                             | ii  |
| CURRICULUM VITAE.....                                                 | iii |
| ÖZGEÇMİŞ.....                                                         | iv  |
| ABSTRACT.....                                                         | v   |
| ÖZET.....                                                             | vii |
| ACKNOWLEDGMENT.....                                                   | ix  |
| ÖNSÖZ.....                                                            | x   |
| TABLE OF CONTENTS.....                                                | xi  |
| LIST OF TABLES.....                                                   | xv  |
| LIST OF ABBREVIATIONS.....                                            | xvi |
| CHAPTER I: INTRODUCTION.....                                          | 1   |
| 1.1. Background to the Study.....                                     | 1   |
| 1.2. Significance of the Study.....                                   | 5   |
| 1.3. Purpose of the Study.....                                        | 6   |
| 1.4. Delimitations of the Study.....                                  | 7   |
| CHAPTER II: LITERATURE REVIEW.....                                    | 8   |
| 2.1. Vocabulary Knowledge in Second Language Acquisition (SLA).....   | 8   |
| 2.2. Aspects of L2 Vocabulary Knowledge.....                          | 9   |
| 2.3. Intentional vs. Incidental Vocabulary Learning.....              | 11  |
| 2.3.1. L2 Vocabulary Learning through Written Input.....              | 13  |
| 2.3.2. L2 Vocabulary Learning through Audio or Audiovisual Input..... | 21  |
| 2.4. Depth of Processing.....                                         | 25  |
| 2.4.1. The Multistore Model.....                                      | 25  |
| 2.4.2. Levels of Processing.....                                      | 26  |

|                                                                           |           |
|---------------------------------------------------------------------------|-----------|
| 2.4.3. Depth of Processing and L2 Vocabulary Acquisition.....             | 27        |
| 2.5. The Involvement Load Hypothesis.....                                 | 28        |
| 2.5.1. Components of the Involvement Load Hypothesis.....                 | 29        |
| 2.5.2. Empirical Evidence for the Involvement Load Hypothesis.....        | 31        |
| 2.5.3. The Link between Input Type and Task-Induced Involvement Load..... | 37        |
| 2.6. Research Questions.....                                              | 39        |
| <b>CHAPTER III: METHODOLOGY.....</b>                                      | <b>41</b> |
| 3.1. Research Design.....                                                 | 41        |
| 3.2. Participants.....                                                    | 41        |
| 3.3. Data Collection Instruments.....                                     | 42        |
| 3.3.1. Video Podcasts.....                                                | 43        |
| 3.3.2. Reading Passages.....                                              | 43        |
| 3.3.3. Target Words.....                                                  | 44        |
| 3.3.4. Vocabulary Tasks.....                                              | 44        |
| 3.3.4.1. Reading-only (RO) Tasks.....                                     | 46        |
| 3.3.4.2. Watching-only (WO) Tasks.....                                    | 46        |
| 3.3.4.3. Fill-in-by-reading (FR) Tasks.....                               | 47        |
| 3.3.4.4. Fill-in-by-watching (FW) Tasks.....                              | 48        |
| 3.3.4.5. Sentence-writing-by-reading (SR) Tasks.....                      | 48        |
| 3.3.4.6. Sentence-writing-by-watching (SW) Tasks.....                     | 50        |
| 3.3.5. Vocabulary Tests.....                                              | 50        |
| 3.4. Data Collection Procedure.....                                       | 52        |
| 3.5. Data Analysis Procedure.....                                         | 54        |
| <b>CHAPTER IV: RESULTS.....</b>                                           | <b>56</b> |
| 4.1. Test of Normality.....                                               | 56        |
| 4.2. Results for Lower-Intermediate Proficiency Level.....                | 56        |

|                                                                                   |           |
|-----------------------------------------------------------------------------------|-----------|
| 4.2.1. Vocabulary Gain from Tasks with Low vs. High Involvement Load.....         | 56        |
| 4.2.1.1. Gain in Productive Word Knowledge.....                                   | 57        |
| 4.2.1.2. Retention in Productive Word Knowledge.....                              | 58        |
| 4.2.1.3. Changes in Productive Word Knowledge within Groups.....                  | 59        |
| 4.2.1.4. Gain in Receptive Word Knowledge.....                                    | 60        |
| 4.2.1.5. Retention in Receptive Word Knowledge.....                               | 61        |
| 4.2.1.6. Changes in Receptive Word Knowledge within Groups.....                   | 62        |
| 4.2.2. Vocabulary Gain from Tasks with Written vs. Audiovisual Input Mode.....    | 62        |
| 4.2.2.1. Gain and Retention in Productive Word Knowledge.....                     | 62        |
| 4.2.2.2. Gain and Retention in Receptive Word Knowledge.....                      | 64        |
| 4.3. Results for Upper-Intermediate Proficiency Level.....                        | 65        |
| 4.3.1. Vocabulary Gain from Tasks with Low vs. High Involvement Load.....         | 65        |
| 4.3.1.1. Gain in Productive Word Knowledge.....                                   | 66        |
| 4.3.1.2. Retention in Productive Word Knowledge.....                              | 67        |
| 4.3.1.3. Changes in Productive Word Knowledge within Groups.....                  | 68        |
| 4.3.1.4. Gain in Receptive Word Knowledge.....                                    | 68        |
| 4.3.1.5. Retention in Receptive Word Knowledge.....                               | 70        |
| 4.3.1.6. Changes in Receptive Word Knowledge within Groups.....                   | 71        |
| 4.3.2. Vocabulary Gain from Tasks with Written vs. Audiovisual Input Mode.....    | 71        |
| 4.3.2.1. Gain and Retention in Productive Word Knowledge.....                     | 72        |
| 4.3.2.2. Gain and Retention in Receptive Word Knowledge.....                      | 73        |
| 4.4. Summary of the Results.....                                                  | 74        |
| <b>CHAPTER V: DISCUSSION AND CONCLUSION.....</b>                                  | <b>75</b> |
| 5.1. Discussion of the Findings.....                                              | 75        |
| 5.1.1. Vocabulary Gain from Various Loaded Tasks at Lower-Intermediate Level..... | 75        |

|                                                                                       |            |
|---------------------------------------------------------------------------------------|------------|
| 5.1.2. Vocabulary Gain from Different Modes of Input at Lower-Intermediate Level..... | 78         |
| 5.1.3. Vocabulary Gain from Various Loaded Tasks at Upper-Intermediate Level.....     | 82         |
| 5.1.4. Vocabulary Gain from Different Modes of Input at Upper-Intermediate Level..... | 85         |
| 5.2. Summary of the Findings and Discussion.....                                      | 89         |
| 5.3. Pedagogical Implications.....                                                    | 90         |
| 5.4. Limitations and Suggestions for Further Research.....                            | 94         |
| <b>REFERENCES.....</b>                                                                | <b>97</b>  |
| <b>APPENDICES.....</b>                                                                | <b>106</b> |
| APPENDIX A Reading-only (RO) Tasks.....                                               | 107        |
| APPENDIX B Watching-only (WO) Tasks.....                                              | 112        |
| APPENDIX C Fill-in-by-reading (FR) Tasks.....                                         | 114        |
| APPENDIX D Fill-in-by-watching (FW) Tasks.....                                        | 119        |
| APPENDIX E Sentence-writing-by-reading (SR) Tasks.....                                | 123        |
| APPENDIX F Sentence-writing-by-watching (SW) Tasks.....                               | 129        |
| APPENDIX G Vocabulary Pre-tests.....                                                  | 133        |
| APPENDIX H Productive Vocabulary Post-tests.....                                      | 135        |
| APPENDIX I Receptive Vocabulary Post-tests.....                                       | 136        |
| APPENDIX J Lexical Production Scoring Protocol.....                                   | 137        |
| APPENDIX K Normality Test Results.....                                                | 138        |
| APPENDIX L Changes in Word Knowledge (Lower-Intermediate Learners).....               | 140        |
| APPENDIX M Changes in Word Knowledge (Upper-Intermediate Learners).....               | 141        |

## LIST OF TABLES

|             |                                                                                   |     |
|-------------|-----------------------------------------------------------------------------------|-----|
| Table 2.1.  | Dimensions of L2 Vocabulary.....                                                  | 10  |
| Table 2.2.  | Task-Induced Involvement Indexes for Various Tasks.....                           | 31  |
| Table 3.1.  | Distribution of Participants.....                                                 | 42  |
| Table 3.2.  | Overview of Tasks and Their Indexes.....                                          | 45  |
| Table 4.1.  | Productive Word Gain and Retention at Lower-Intermediate Level (Index Based)..... | 57  |
| Table 4.2.  | Initial Gain in Productive Word Knowledge at Lower-Intermediate Level.....        | 57  |
| Table 4.3.  | Retention in Productive Word Knowledge at Lower-Intermediate Level.....           | 58  |
| Table 4.4.  | Receptive Word Gain and Retention at Lower-Intermediate Level (Index Based).....  | 60  |
| Table 4.5.  | Initial Gain in Receptive Word Knowledge at Lower-Intermediate Level.....         | 60  |
| Table 4.6.  | Retention in Receptive Word Knowledge at Lower-Intermediate Level.....            | 61  |
| Table 4.7.  | Productive Word Gain and Retention at Lower-Intermediate Level (Input Based)..... | 63  |
| Table 4.8.  | Receptive Word Gain and Retention at Lower-Intermediate Level (Input Based).....  | 64  |
| Table 4.9.  | Productive Word Gain and Retention at Upper-Intermediate Level (Index Based)..... | 66  |
| Table 4.10. | Initial Gain in Productive Word Knowledge at Upper-Intermediate Level.....        | 66  |
| Table 4.11. | Retention in Productive Word Knowledge at Upper-Intermediate Level.....           | 67  |
| Table 4.12. | Receptive Word Gain and Retention at Upper-Intermediate Level (Index Based).....  | 69  |
| Table 4.13. | Initial Gain in Receptive Word Knowledge at Upper-Intermediate Level.....         | 69  |
| Table 4.14. | Retention in Receptive Word Knowledge at Upper-Intermediate Level.....            | 70  |
| Table 4.15. | Productive Word Gain and Retention at Upper-Intermediate Level (Input Based)..... | 72  |
| Table 4.16. | Receptive Word Gain and Retention at Upper-Intermediate Level (Input Based).....  | 73  |
| Table 4.17. | Summary of the Findings.....                                                      | 74  |
| Table J.    | Lexical Production Scoring Protocol.....                                          | 137 |
| Table K1.   | Normality Test of the Lower-Intermediate Participants.....                        | 138 |
| Table K2.   | Normality Test of the Upper-Intermediate Participants.....                        | 139 |
| Table L1.   | Changes in Productive Word Knowledge (Lower-Intermediate).....                    | 140 |
| Table L2.   | Changes in Receptive Word Knowledge (Lower-Intermediate).....                     | 140 |
| Table M1.   | Changes in Productive Word Knowledge (Upper-Intermediate).....                    | 141 |
| Table M2.   | Changes in Receptive Word Knowledge (Upper-Intermediate).....                     | 141 |

## **LIST OF ABBREVIATIONS**

|        |                                         |
|--------|-----------------------------------------|
| EFL:   | English as a Foreign Language           |
| ESL:   | English as a Second Language            |
| ESP:   | English for Specific Purposes           |
| FR:    | Fill-in-by-reading Task                 |
| FW:    | Fill-in-by-watching Task                |
| IEP:   | Intensive English Program               |
| ILH:   | Involvement Load Hypothesis             |
| L1:    | First Language (Native Language)        |
| L2:    | Second Language                         |
| LTS:   | Long-term Store (Long-term Memory)      |
| RO:    | Reading-only Task                       |
| SLA:   | Second Language Acquisition             |
| SPSS:  | Statistical Package for Social Sciences |
| SR:    | Sentence-writing-by-reading Task        |
| STS:   | Short-term Store (Short-term Memory)    |
| SW:    | Sentence-writing-by-watching Task       |
| TOEFL: | Test of English as a Foreign Language   |
| VKS:   | Vocabulary Knowledge Scale              |
| WO:    | Watching-only Task                      |

## **CHAPTER I: INTRODUCTION**

This chapter provides the background to the present study, its significance in the field and its purpose with the research questions addressed prior to the study.

### **1.1. Background to the Study**

Vocabulary is one of the most important elements of foreign or second language acquisition. The more knowledge of words a second language (L2) learner acquires, the easier his or her comprehension will be. Otherwise, obtaining a limited knowledge of vocabulary in the target language will hinder effective communication. Therefore, learning a great deal of vocabulary has a facilitating role not only in the comprehension but also in the production of the language. To emphasize how vocabulary knowledge and language acquisition are interconnected, Schmitt (2000) stated that “lexical knowledge is central to communicative competence and to the acquisition of a second language” (p. 55). Nation (2001) also regarded vocabulary knowledge and language use as complementary. This means that knowledge of words positively affects language use, and language use results in an increase in vocabulary size. Folse (2006) further highlighted that acquiring a vast amount of vocabulary is essential for L2 learners to perform better in four major language skills: listening, speaking, reading and writing. In short, having a large body of vocabulary plays a crucial role in helping L2 learners express themselves more clearly and understand the messages coming from their interlocutors much better.

Research on second language vocabulary acquisition has revealed that L2 learners need to know at least 3,000 word families to understand spoken texts (Webb and Rodgers, 2009) and 8,000 word families to understand written texts (Nation, 2006). These two numbers only represent 95% lexical coverage for spoken texts and 98% lexical coverage for written texts. Even such large figures are not sufficient to reach 100% lexical coverage. As a result, it is obvious that learning vocabulary is one of the biggest obstacles that L2 learners have to overcome in order to be successful language users.

For the acquisition of L2 vocabulary, two central views have been formed. One of them is intentional learning which is described as “the deliberate committing to memory of

thousands of words (their meaning, sound, and spelling)” (Hulstijn, 2003, p. 349). The other term is incidental learning which is defined as “picking up” words while engaging in a communicative task (e.g., listening or reading) with a focus on meaning. In this sense, incidental learning takes place subconsciously (Krashen, 1989) or it happens as the by-product of an intention that is initially directed at learning something else (Schmidt, 1994). Compared to intentional learning, incidental learning has been regarded as the major source of L2 vocabulary knowledge (Nation, 2013). This is mainly because native speakers also acquire most of their lexicon through daily activities in incidental settings before intentionally learning their mother tongue at school. Because incidental vocabulary learning involves guessing the meaning of unknown words by using the contextual clues, it requires learners to be fully involved while inferring the meaning, which results in deeper cognitive processing and better retention. In contrast, intentional vocabulary learning (e.g., word substitution, crossword puzzles, etc.) is similar to rote learning. Meaning can be acquired without undergoing any mental processes, but very few words learned intentionally can be retained (Nagy and Herman, 1987). For communicative competence, language learners need to acquire large quantities of words, which cannot be achieved only through intentional vocabulary activities. That is why incidental learning receives great attention in L2 vocabulary acquisition.

Reading has been considered to be a primary resource for incidental vocabulary learning (Brown, Waring and Donkaewbua, 2008; Horst, Cobb and Meara, 1998; Huckin and Coady, 1999; Krashen 1993; Kweon and Kim, 2008; Pigada and Schmitt, 2006; Waring and Takaki, 2003). However, incidental vocabulary learning through reading has a rather low pick-up rate. For example, Waring and Takaki (2003) found that only one word can be acquired from one hour of reading. Moreover, Horst, Cobb and Meara (1998) suggested that L2 learners need to encounter a new word about 8 times to be able to acquire it. Thus, it is predominantly suggested that learning words through reading alone is a rather time-consuming process, and it is susceptible to errors (Peters, Hulstijn, Sercu and Lutjeharms, 2009). Therefore, incidental vocabulary acquisition through mere reading may not be adequate.

Research has also focused on listening as a medium for incidental vocabulary acquisition (Fatalaki, 2014; Lin, 2010; Jing and Jianbin, 2009; van Zeeland and Schmitt, 2013; Vidal, 2011). Studies comparing the benefits of listening to reading in incidental vocabulary learning have found that reading initially leads to more word gains than listening, but vocabulary knowledge gained from listening seems to be retained better. However, Vidal (2011) stated that more encounters with the new words are required in listening (5-6 times) than reading (2-3 times) for more durable effects. Van Zeeland and Schmitt (2013) further acknowledged that knowledge of meaning is more difficult to be acquired through listening. The relatively small vocabulary gain from listening input is found to be moderated by a wide range of factors such as the difficulty of real-time listening, the limited capacity of the working memory and the challenge of L2 lexical recognition and segmentation in spoken discourse. Considering these problems associated with listening, it is reasonable to consult to other types of input for incidental vocabulary learning.

When reading or listening alone appears not to be a particularly efficient resource for expanding L2 word knowledge, Paribakht and Wesche (1997) noted that L2 learners cannot be expected to learn a considerable amount of vocabulary by themselves without any instructional intervention due to the unpredictable nature of incidental learning. Subsequently, Laufer (2003) suggested that this instructional guidance can be achieved through meaning-focused vocabulary tasks that would direct learners' attention to the form and meaning of specific vocabulary items chosen by the teacher. This might help L2 learners acquire and retrieve these target words better. Here, a pertinent question regarding the effectiveness of certain vocabulary tasks over others has emerged. To solve this issue, Laufer and Hulstijn (2001) proposed the Task-induced Involvement Load Hypothesis (ILH) to determine the efficacy of vocabulary tasks in fostering L2 vocabulary knowledge. Their hypothesis stated that a vocabulary task can be effective if it induces more learner involvement by incorporating higher degrees of the three components of the hypothesis (i.e., need, search and evaluation). This means that if learners perform a task that requires them to deal with the target words by answering some comprehension questions relevant to

the words, looking up their definitions, filling the gaps with them or using them to write some sentences, the likelihood of recalling these words later will be higher.

Based on the hypothesis, many studies have aimed at proving that supplementing reading or listening texts with word-focused vocabulary tasks under incidental conditions yields more vocabulary gain and retention (Fatalaki, 2014; Hulstijn and Laufer, 2001; Jing and Jianbin, 2009; Keating, 2008; Kim, 2011; Lu, 2013; Mármol and Sánchez-Lafuente, 2013; Teng, 2015; Tu, 2004; Yaqubi, Rayati and Gorgi, 2010). The findings of these studies indicated that engaging learners in a reading or listening activity complemented with an involving vocabulary task (e.g., gap filling and sentence/composition writing) is more effective in enhancing L2 lexical knowledge than asking them to read or listen to a text alone. Consequently, the findings of previous research proved that additional vocabulary tasks can be beneficial in facilitating incidental vocabulary learning.

Previous research on the link between additional vocabulary tasks and incidental vocabulary learning has fundamentally concentrated on utilizing written texts (reading as a medium) or audio materials (listening as a medium) for the design of such tasks. However, viewing (i.e., the audiovisual medium) is also found to be a productive source of incidental L2 vocabulary uptake (Marefat and Hassanzadeh, 2014; Neuman and Koskinen, 1992; Rodgers, 2013; Sydorenko, 2010). For example, Neuman and Koskinen (1992) found that L2 learners acquired more lexical knowledge when they watched an educational television program (with or without L2 captions) compared to when they read its transcript while listening to the same program or when they only read the transcript. They concluded that verbal and pictorial clues in audiovisual materials are found to be useful in assisting L2 learners infer the meaning of unknown words. Similarly, Marefat and Hassanzadeh (2014) compared the effects of different types of input (written, aural and audiovisual) on incidental L2 lexical acquisition. They found that while there was no difference between the learners who read four texts or listened to four podcasts on the same topic, the learners who watched four vodcasts significantly outperformed them on both immediate and delayed post-tests.

The aforementioned studies on the audiovisual medium have not explored the potential contribution of additional vocabulary tasks to incidental vocabulary learning. To the knowledge of the researcher, there has been merely one study that looked into the efficacy of vocabulary tasks that are generated from the assumptions of the Involvement Load Hypothesis and implemented through the audiovisual input. In Hassanzadeh's (2016) study, it was found that highly involving tasks that incorporate an audiovisual resource result in more vocabulary gain than those that rely on a written material. However, this finding needs to be validated with more empirical research in the field.

## **1.2. Significance of the Study**

As stated before, a good deal of vocabulary knowledge is a prerequisite for successful second language acquisition (Coady and Huckin, 1997; Nation, 2001; Nunan, 1999; Schmitt, 2000). However, the very small amount of vocabulary uptake through reading or listening alone has led several researchers to divert their attention to using additional vocabulary tasks for incidental vocabulary learning (Hulstijn and Laufer, 2001; Laufer, 2003; Wesche and Paribakht, 2000). With a wide choice of tasks in the field, the need to assess their potential benefits in promoting L2 vocabulary knowledge has emerged. As a result, Laufer and Hulstijn (2001) proposed the Task-induced Involvement Load Hypothesis to help teachers and researchers evaluate task efficacy, and a large body of research testing the hypothesis have centred on comparing a number of vocabulary tasks consisting of various degrees of task-induced involvement load (Fatalaki, 2014; Hulstijn and Laufer, 2001; Jing and Jianbin, 2009; Keating, 2008; Kim, 2011; Lu, 2013; Mármol and Sánchez-Lafuente, 2013; Teng, 2015; Tu, 2004; Yaqubi, Rayati and Gorgi, 2010).

However, most of the aforementioned studies have primarily benefitted from written or audio materials, measured only passive word recall and lacked diverse participant groups at different levels of language competence. Furthermore, the hypothesis seems to be relatively unexplored in Turkey since only one study exclusively tested its assumptions by implementing four different vocabulary tasks (fill-in and rewriting) through two reading texts with the target words glossed or looked up (Karalık, 2016). Still, Karalık (2016) also designed his tasks only with the written input. All in all, a review of the previous studies

pointed out that several issues such as input diversity, involvement load equivalency, knowledge type and L2 proficiency have been areas of investigation neglected by researchers. Therefore, the present study set out to fill this gap in the literature by investigating the effects of task-induced involvement load, input type (or modality) and language proficiency on Turkish EFL (English as a Foreign Language) learners' receptive and productive word knowledge.

### **1.3. Purpose of the Study**

To fill the addressed gaps in the field, the present study aimed to provide empirical evidence to the preliminary research on the ILH (Laufer and Hulstijn, 2001) by exploring task efficacy along with input type, knowledge type and L2 proficiency. In this way, the present study aimed to determine the predictive value of certain tasks in fostering incidental L2 vocabulary knowledge and thus pedagogically help language teachers and material developers design tasks that best enhance incidental vocabulary learning. Therefore, the research questions addressed in this study are as follows:

- 1) Do vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at lower-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 2) Do vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at lower-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 3) Do vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at upper-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?

- 4) Do vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at upper-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?

#### **1.4. Delimitations of the Study**

The present study has several delimitations. In terms of population, it is only limited to a particular group of university students studying at an Intensive English Program (IEP) at a preparatory school of a state university in Turkey. Therefore, the results cannot be generalized to other groups of foreign language learners. In terms of the time of the research implementation, it is only restricted to 2016-2017 academic year. Finally, the present study only measures the participants' vocabulary knowledge and gain of the target words. It does not aim to assess any other language skill.

## **CHAPTER II: LITERATURE REVIEW**

This chapter provides a brief definition for the fundamental concepts associated with L2 vocabulary acquisition. It then gives a general background to the pioneering studies conducted in the field of incidental vocabulary learning. It also explains the theoretical framework of the present study and describes the empirical research that relates to it. Finally, it concludes with the purpose and research questions of the present study.

### **2.1. Vocabulary Knowledge in Second Language Acquisition (SLA)**

Knowing a sufficient amount of vocabulary has been viewed as a significant determinant that influences L2 learners' ability to comprehend and express themselves clearly in English. Considering an educated adult native English speaker whose lexical knowledge can reach up to 20,000 word families (Goulden, Nation and Read, 1990), it is no wonder that L2 learners are rather apprehensive about their lack of L2 vocabulary. When it comes to the number of words to be acquired by L2 learners, Laufer (1997) and Nation (1990) stated that 5,000 words constitute the most basic lexical set necessary to understand general, non-specialized or unsimplified texts in English. Schmitt (2000) similarly pointed out that L2 learners need to know approximately 2,000 words to sustain conversations, 3,000 word families to comprehend authentic reading texts and around 10,000 words to understand complex academic passages. According to Nunan (1999), having a vast amount of vocabulary knowledge is believed to enable learners to "outperform their competence" (p. 103). This means that learners would be able to deal with unpredictable situations that emerge during communication to a certain extent. Laufer (2003) also stated that without an adequate supply of vocabulary, it would be unlikely for speakers to transmit their messages or interact with each other in the relevant language. Therefore, it is obvious that L2 learners are required to master an extensive stock of words to function well in English.

A substantial bulk of research in L2 vocabulary acquisition has highlighted the importance of vocabulary development in reading comprehension (Haynes and Baker, 1993; Henriksen, Albrechtsen and Haastrup, 2004; Hu and Nation, 2000; Laufer, 1992; Schmitt, Jiang and Grabe, 2011; Stæhr, 2008; Tannenbaum, Torgesen and Wagner, 2006), in oral

communication (Hincks, 2003; Joe, 1998; Meara, 1996), in listening comprehension (Bonk, 2000; Ellis, 1994; Goh, 2000; Milton, Wade and Hopkins, 2010; van Zeeland and Schmitt, 2012) and in written performance (Albrechtsen, Haastrup and Henriksen, 2008; Hinkel, 2001; Laufer and Nation, 1995). Thus, it is commonly accepted by many scholars that lexical competence is at the very core of communicative competence (Coady and Huckin, 1997). For instance, L2 learners should know 98-99% of the words in a written text (about one unknown word in every 50-100 following words) to understand 70% of authentic reading texts (Hirsh and Nation, 1992; Hu and Nation, 2000; Nation, 2001). Likewise, 95% lexical coverage (the ratio of known and unknown words in the text) is recommended to largely comprehend narrative spoken texts (van Zeeland and Schmitt, 2012; Webb and Rodgers, 2009).

On the other hand, the prevailing view among most researchers is that it is not possible for L2 learners to memorize the vast majority of words in the English lexicon because words in English usually convey multiple meanings due to their linguistic, psycholinguistic and sociolinguistic features (Zimmerman, 1997). This implies that lexical competence ultimately turns into an ability that is beyond explaining the definition of a word. It involves knowing such components of a word as its general frequency, grammatical and semantic usage, collocations, derivations, etc. As a result, attempts at discovering the most productive methods for the teaching of vocabulary have originated from the objective that the process of lexical acquisition should be facilitated for L2 learners. Therefore, it is no surprise that a great deal of research in L2 vocabulary teaching have primarily investigated the factors that contribute to the expansion of L2 word knowledge in the most effective ways.

## **2.2. Aspects of L2 Vocabulary Knowledge**

Before reviewing the research on L2 vocabulary acquisition, it is crucial to describe what it means to know a 'word' in a foreign language. According to Nation (1990, 2001), knowing a word is basically defined as understanding the meaning of a lexical item or providing the correct form for a specific meaning. However, defining lexical knowledge is not as simple as that because the process of acquiring a new word involves several other aspects such as

the grammatical position, collocations, frequency, appropriateness and associations of the word to be learned. Nation (1990) categorised these aspects under two dimensions of word knowledge: receptive and productive (see Table 2.1). Receptive word knowledge means the ability to understand or recognise different aspects of a word in the language input (e.g., its meaning, spoken/written form and appropriateness), whereas productive word knowledge is defined as the ability to recall and produce the correct word (its spoken/written form, grammatical features, synonyms, collocations, etc.) to convey a particular meaning.

**Table 2.1. Dimensions of L2 Vocabulary**

|                |                       |                                                           |
|----------------|-----------------------|-----------------------------------------------------------|
| <b>Form</b>    | Spoken                | R What does the word sound like?                          |
|                |                       | P How is the word pronounced?                             |
|                | Written               | R What does the word look like?                           |
|                |                       | P How is the word written and spelled?                    |
|                | Word parts            | R What parts are recognisable in this word?               |
|                |                       | P What words parts are needed to express meaning?         |
| <b>Meaning</b> | Form and meaning      | R What meaning does this word form signal?                |
|                |                       | P What word form can be used to express this meaning?     |
|                | Concept and referents | R What is included in the concept?                        |
|                |                       | P What items can the concept refer to?                    |
|                | Associations          | R What other words does this word make us think of?       |
|                |                       | P What other words could we use instead of this one?      |
| <b>Use</b>     | Grammatical functions | R In what patterns does the word occur?                   |
|                |                       | P In what patterns must we use this word?                 |
|                | Collocations          | R What words or types of words occur with this one?       |
|                |                       | P What words or types of words must we use with this one? |
|                | Constraints on use    | R Where, when and how often would we meet this word?      |
|                |                       | P Where, when and how often can we use this word?         |

**Note.** R: Receptive and P: Productive

**Source.** Nation (2001, p. 27)

Although Nation (1990) tried to create a framework to explain in detail what L2 vocabulary knowledge comprises, the distinction between receptive word knowledge and productive word knowledge has been problematic when it comes to measuring lexical gains in

vocabulary learning studies. Assessing productive word knowledge is a complicated and laborious process because data about a learner's productive word knowledge can be gathered through that learner's actual production of the target language. For that, researchers or teachers have to record the learner's oral or written language input during performance and analyze it by means of a specifically designed scale, instrument, scoring protocol, etc. As a result, learners' receptive word knowledge has been mostly favoured as an indication of their vocabulary gains in the SLA research.

To report a learner's receptive and productive word knowledge accurately, it is also fundamental to determine if this knowledge is 'partial' or 'precise' (Henriksen, 1999). Partial word knowledge indicates that the learner vaguely recognises the meaning or form of a word, while precise word knowledge means that the learner has full access to the meaning as well as the form. These constructs should be regarded as the two ends of a continuum since a learner may first recognise the form of a word and then retrieve the meaning after several encounters. This is because learners need full knowledge to recall a word, while a small amount is enough to recognise a word among several options (Norman, 1976). Therefore, only testing receptive word knowledge does not necessarily help deduce the conclusion that new words will be stored in the learner's lexicon. Thus, SLA research studies on vocabulary learning should also focus on promoting and assessing productive knowledge of unknown words. Based on this suggestion, the present study measured both receptive (operationalised as the recall of the word meaning for a given form) and productive (operationalised as the recall of the word form for a given meaning) vocabulary gains through partial and precise representation of word knowledge.

### **2.3. Intentional vs. Incidental Vocabulary Learning**

L2 vocabulary teaching and learning practices have been carried out under the influence of two approaches to SLA vocabulary acquisition: intentional learning and incidental learning. Intentional vocabulary learning occurs when learners make a conscious effort to learn new words and commit them to their long-term memories by using specific memorization techniques such as rehearsals usually to get prepared for an upcoming test (Gass, 1999; Hulstijn, 2003). In contrast, incidental vocabulary learning has been referred to as learning

new words not because of paying deliberate attention to them but because of engaging in a communicative task with a focus primarily on meaning (Hulstijn, 2001; Nation, 2001). However, as Hulstijn and Laufer (2001) noted, incidental learning is not necessarily equal to unattended learning. Similarly, Schmidt (1994) also acknowledged that for incidental learning to take place, learners are required to be conscious to some extent when they notice the new words that occur in the input. It is thus problematic to distinguish the two kinds of learning by means of the absence or presence of attention only.

A clear-cut distinction between intentional and incidental learning is difficult to reach because the kind and nature of learner attention also determine the acquisition of new items. For instance, Craik and Lockhart (1972) noted that conscious attention is conditional upon how deeply learners engage with a new word and whether they can grasp an insightful understanding of the word from that encounter. Whether this level of attention is focal or peripheral is another factor to consider when learners try to learn new words in different situations (R. Ellis, 1999). In intentional learning, learners concentrate on form or form-meaning associations more than comprehending the message content. In other words, they deliberately place their focal attention on the linguistic form. On the other hand, in incidental learning comprehension is the key, so this focal attention is directed to meaning, whereas peripheral attention is placed on form as well. These implicit and explicit cognitive processes clearly show that incidental learning is not purely unconscious or without any attention.

It is obvious that learner attention cannot entirely account for the difference between intentional and incidental learning. Hence, Gass (1999) convincingly argued that such attention-based approaches to learning are blurred because it is not possible to access to learners' minds to see if they are actively paying attention to target words. Therefore, Hulstijn (2001) operationalised incidental learning as happening when learners learn unfamiliar words as the by-product of a communicative task without being aware of an upcoming vocabulary test designed to check their lexical gains or retention. This definition of incidental learning is used in the present study to describe the learning process that the

participants manifested through various tasks without being forewarned of any vocabulary tests.

### **2.3.1. L2 Vocabulary Learning through Written Input**

In the first language (L1) acquisition research, it is well-known that reading is a considerably effective means of building vocabulary (Sternberg, 1987). Reading is also important for learning a second or foreign language. As Krashen suggested (1989), reading extensively in a foreign language can be beneficial for vocabulary development in addition to the improvement of grammar, spelling and writing skills. For a number of reasons, extensive reading has been considered powerful by many researchers (Horst, Cobb and Meara, 1998; Huckin and Coady, 1999; Krashen 1993; Wodinsky and Nation, 1988). For one thing, it exposes learners to substantial amounts of material in a contextualized fashion, which is often enjoyable and motivating (Grabe and Stoller, 2002). It is also pedagogically efficient (Huckin and Coady, 1999) because learners are engaged in two activities (reading and vocabulary learning) simultaneously and often autonomously, which eventually boosts reading fluency and facilitates vocabulary retention. Nation (2001) also agreed that extensive reading can be a practical solution for language development outside the classroom particularly in EFL contexts.

Considering all the reasons above, a number of studies conducted over the last few decades have centred on the causal link between extensive reading and vocabulary development in the ESL (English as a second language) or EFL research. For example, Horst, Cobb and Meara (1998) examined the extent to which new words can be learned as a result of extensive reading. The study involved 34 Arabic low-intermediate EFL learners enrolled in an intensive English program at a university. To ensure that the learners would be exposed to the whole book, the teacher read aloud a graded reader in the classroom over a period of six sessions while the learners listened to it. 45 target words were selected and grouped under six bands (ranging from 2 to 7 occurrences) based on their in-text frequency. A week before and after the sessions, the learners took two tests: a multiple-choice test requiring the recognition of a correct definition and a word-association test assessing the form-meaning link between two words by excluding the third one that does not fit. Nation's (1990)

Vocabulary Levels Test was also administered in the beginning to see if vocabulary size positively affects learners' ultimate lexical gains. The findings indicated that the learners were able to recognise about one in every five unknown words (about 22% of the target words) and pinpoint an average of 16% correct form-meaning associations. The target words appeared eight or more times were acquired most regardless of their frequency in English. It was also found that the learners with a larger vocabulary size tend to learn more words.

In another study, Waring and Takaki (2003) wanted to find out the pick-up rate at which new words can be learned from a graded reader. Their participants included 15 Japanese EFL learners at a lower-intermediate or above level studying at a Japanese university. They selected 25 target words appearing from one to 17 times and then changed them into non-words (e.g., replacing *house* with *windle*). The learners read the book all at once in an hour and were immediately tested through three tests (a word-form recognition test, a meaning-translation test and a multiple-choice recognition test). The tests respectively required the learners to circle the target words given in a different text, supply their L1 equivalents and choose their correct meanings from four options. The delayed post-tests were administered after one week and three months later. The results showed that on the immediate post-test, the learners recognised an average of 61% correct forms, translated 18% of the words correctly and identified 42% correct meanings. However, three months later the learners lost about half of their word knowledge which corresponded to only one word acquired from one hour of reading. Moreover, in-text frequency was found to be an important indicator, but a definite number for the optimum level of encounters with the target words could not be set. This is because the learners recalled about half of the word forms that appeared eight or more times and supplied the correct meaning for 6% of those that appeared more than 15 times after a three-week interval. As a result, Waring and Takaki (2003) concluded that very few new words can be learned and retained through graded reading since learners usually concentrate on their enjoyment and comprehension of the text. However, making such words explicit through highlighting or bolding can help learners notice and learn these words better (Schmidt, 1990).

Pigada and Schmitt (2006) criticised pre-existing reading-based vocabulary studies because they assessed few target words, lasted for a brief time and did not measure any aspects of knowledge except for meaning. Therefore, they designed a case study with a Greek beginning learner of French to look into the effects of extensive reading on full and partial learning of spelling, meaning and grammatical function. At the time, the adult participant was enrolled in a non-linguistic postgraduate program in England. Over a one-month period, he finished reading four graded readers which incorporated 133 target words (nouns together with their articles and verbs along with their prepositions). The in-text frequency of the words ranged from one to 20 times. During the interviews done before and after the reading program, the learner was tested on the target words regarding their forms, meanings and grammatical features. He was asked to write down the correct forms, express the appropriate meanings (or any recognised words) and produce the right determiners and prepositions for the nouns and verbs respectively. The results revealed that the learner achieved 24% word gains in spelling, 13% in meaning and 19% in grammatical behaviour. Any uptake in all three aspects depended on 20 or above encounters with the relevant word. Verb forms and meanings were learned better than their prepositions, while noun spellings and their articles were retained more than their meanings. On average, the pick-up rate was one in every 1.5 words. However, it was a case study with no subsequent post-tests to observe the durability of retention, so the findings are questionable with regard to their generalisability.

In examining vocabulary learning through the reading of authentic materials (three graded readers) by 12 Korean intermediate EFL learners, Kweon and Kim (2008) looked at such factors as word class and in-text frequency. They selected 367 target words of three word classes (verbs, nouns and adjectives) and sorted them into three groups (appearing 1-6, 7-19 and 20 or above times). The learners were assigned a few chapters to read before each class. Then they exchanged their ideas on a book-related theme in the classroom. At the end of each class, a response journal was written based on the chapters studied on that day. On the final day of treatment and 4 weeks later, the post-tests were administered by asking learners to give a *Yes*, *No* or *Not Sure* response regarding the meaning of the target words.

The results of these self-report tests implied that the learners significantly increased their knowledge of the target words from the pre-test to the immediate post-test and retained that knowledge on the delayed post-test too. Nouns were the ones that were learned and recalled more easily followed by adjectives and verbs. Moreover, the learning and retention of more frequent words were greater than the words of lower frequency, but the less frequent ones were also acquired if their meanings were conducive to comprehension. As a result, Kweon and Kim (2008) asserted that learners can also acquire less common words when they focus on meaning upon noticing that such words are crucial to the story. However, they cautioned that their findings should be interpreted carefully since self-report checklists rely highly on self-evaluation which is vulnerable to the overestimation of actual learning.

The research reviewed in this section indicated that incidental vocabulary learning can occur as a result of extensive reading of L2 materials over a prolonged period of time. However, the studies that investigated the benefits of extensive reading arrived at different conclusions concerning the number of new words learned and retained after graded reading. To illustrate, word gains ranged from only one of the 25 new words encountered in a graded reader (Waring and Takaki, 2003) to about one of every 1.5 unknown items chosen from four graded readers (Pigada and Schmitt, 2006). There seems to be a wide range of variables such as the learner vocabulary size, the number of in-text occurrences of the target words, word class and the general frequency of the items in the target language that relatively affect the amount of vocabulary intake. Nevertheless, all these key factors underline the conclusion that learners need to be exposed to new words repeatedly over an extended period of time in order to commit them to their lexicon. Even then there is no guarantee of full lexical mastery (Horst, Cobb and Meara, 1998).

In light of previous research, it is clear that extensive reading is an indispensable source of vocabulary knowledge because it helps learners reinforce the knowledge of the words they already know and build deeper form-meaning associations. Still, very few new words seem to be learned since reading for meaning does not necessarily result in vocabulary growth (Huckin and Coady, 1999). There are several reasons for this small amount of lexical uptake. First of all, L2 vocabulary acquisition through extensive reading seems to be rather

slow because learners need to encounter a new word around 8 times to be able to acquire it (Horst, Cobb and Meara, 1998; Nagy, 1997; Nation and Wang, 1999), which means that they have to devote a huge amount of their time to reading. On the contrary, they may disregard unknown words in the first place since these words might not be critical to the comprehension of the gist (Nation and Coady, 1988; Zahar, Cobb and Spada, 2001). Furthermore, implicit vocabulary instruction in the form of extensive reading is an error-prone process because even if learners deduce the meaning of unfamiliar words, their guessing may not correspond to the correct representations of these words. Lexical acquisition is also an incremental process that starts with exposure to input and progresses by means of practice (Horst, Cobb and Meara, 1998; Nation, 2001; Zahar, Cobb and Spada, 2001). It demands that L2 learners should be provided with sufficient input and this exposure must be reinforced with repeated encounters. As a result, considering the circumstances of formal language learning settings extensive reading sounds an unrealistic approach to vocabulary instruction.

The low rates of vocabulary growth from extensive reading and the challenges of adopting such a cumulative approach in instructed language contexts initiated a search for more efficient vocabulary teaching methods. Conforming to this view, many researchers turned to shorter reading materials complemented with form-focused activities designed to teach L2 vocabulary (Hulstijn and Laufer, 2001; Knight, 1994; Laufer, 2003; Llach, 2009; Min, 2008; Paribakht and Wesche, 1997; Peters, Hulstijn, Sercu and Lutjeharms; 2009). For instance, Paribakht and Wesche (1997, 1999, 2000) conducted several studies to compare the effects of two reading approaches (narrow reading and reading plus text-based vocabulary exercises) on intermediate ESL college students' lexical growth. They implemented a one-semester reading-based program at a Canadian university. While the learners in the Reading Only group read more authentic texts including the target words, the Reading Plus group completed a series of vocabulary activities ranging from recognition to production (e.g., matching definitions/synonyms, supplying the correct part of speech, ordering scrambled words, etc.) instead of doing more reading. According to the measurement of Paribakht and Wesche's (1993) Vocabulary Knowledge Scale (VKS) (a 5-

point scale used to measure perceived and actual word knowledge by getting learners to self-report how familiar they are with the target words), both treatments resulted in word gains, but it was noted that reading comprehension combined with vocabulary exercises was more productive than reading only. The Reading Plus group recognized more words and also used these words productively by writing sentences, while the Reading Only group only gained receptive word knowledge. Paribakht and Wesche's (1997, 1999, 2000) findings inspired further studies to concentrate particularly on word-focused tasks by revealing that form-focused vocabulary activities can lead to more gains in L2 vocabulary knowledge than reading multiple texts.

In a similar line of research, Min (2008) also explored the efficacy of word-focused activities over reading-only tasks. 50 Chinese intermediate EFL high school students from Taiwan participated in the study and underwent two treatments. Over a span of 4 weeks, the learners in the reading plus vocabulary enhancement (RV) group received several reading comprehension texts with the 50 boldfaced target words and did many vocabulary exercises (e.g., matching definitions, gap filling, scrambled sentences, etc.) triggering both receptive and productive word knowledge. In contrast, the narrow reading (NR) group did not complete any vocabulary enhancement activities, but they read the same materials along with further topic-related supplementary texts. A modified version of the VKS was administered to the learners to measure their acquired and retained knowledge of the target items. The results demonstrated that both groups gained significantly more target words after the treatment and 3 months later, but the RV group significantly outperformed the NR group in their receptive and productive word knowledge on both post-tests. The significant improvement in the NR group was explained by several reasons. One reason might be the highlighted target words that might have directed learner attention to these items (Schmidt, 1990), which would not normally happen in real-world reading. Some learners were also able to write meaningful and grammatical sentences despite not getting any explicit vocabulary instruction. This implied that such learners were intrinsically motivated enough to employ their own strategies to guess the meaning of the target words.

Peters, Hulstijn, Sercu and Lutjeharms (2009) further investigated the effectiveness of vocabulary tasks with 137 Dutch highly proficient learners of German enrolled in a college in Belgium. Under two enhancement conditions (prior announcement of a vocabulary task or a vocabulary test), the learners were divided into four groups: incidental only (a reading comprehension task), incidental plus (a reading comprehension + vocabulary task), intentional only (a reading comprehension + knowledge of vocabulary tests) and intentional plus (a reading comprehension + vocabulary task + knowledge of vocabulary tests). The learners all worked on computers. By clicking, they could look up the meaning of 16 substitute words while reading a text and during the vocabulary task that demanded an L1/L2 explanation for the words. Shortly after and 2 weeks later, all of them took a word-form recognition test (saying yes/no if the given words appeared before), an isolation recall test (defining the given words) and a context recall test (defining the boldfaced words in the sentences extracted from the text). The frequency of learners' look-up behaviour was measured, and it was seen that they looked up significantly more words and used the online dictionary more often when forewarned of a vocabulary test. On the recognition test, the plus groups reported having seen significantly more words than the reading-only groups. The same pattern was maintained on the recall tests favouring the groups who completed the vocabulary task. Interestingly, although vocabulary test announcement positively affected recognition, it did not prompt more retention. Instead, word relevance (whether the target words are relevant or not to text comprehension questions) and vocabulary task performance were found to have a larger impact because they promoted elaborate processing and strengthened the form-meaning links through repetition of the target words following reading.

In another study carried out by Llach (2009) with 27 German upper-intermediate high school learners of Spanish, three methods of vocabulary instruction were compared: reading only, glossed reading comprehension and sentence writing. The learners assigned to the reading-only task read a text and answered three multiple-choice comprehension questions requiring no knowledge of the 10 target words. The reading comprehension task with the glossed target words in German asked the learners to read the same text and

answer ten multiple-choice comprehension questions into which the target words were incorporated. It was considered as a FonF (focus on form) activity because it directed the learners' attention to the form-meaning link in a communicative situation. As for the third group, they wrote ten sentences with the target words presented in a bilingual list with their L1 definitions. This task was regarded as a FonFs (focus on forms) activity because the target words were explicitly studied in an isolated and decontextualized manner without the inclusion of a communicative activity. The results of a receptive (meaning recall) and a productive (form recall) immediate post-test showed that the FonFs group significantly outscored the FonF group which in turn had significantly higher scores than the reading-only group although the differences on the meaning recall test showed little variance. The groups also achieved parallel scores after a three-week interval, but this time the difference was found to be significant only between the reading-only and FonFs groups in their receptive word knowledge. However, the FonFs group also lost most of their knowledge over time. Overall, the findings pointed to the reading-only task being the least effective condition while suggesting that explicit focus on form-meaning associations positively affects the learning of lexical items especially in the short term. It was suggested that durable and solid form-meaning connections can only be formed through practice and repetition.

Empirical evidence that the above-mentioned studies put forward supported the superiority of direct vocabulary instruction through word-focused tasks and vocabulary enhancement techniques over extensive reading. Therefore, the need for explicit focus on L2 vocabulary under implicit conditions has been acknowledged by several researchers (Hulstijn, Hollander and Greidanus, 1996; Laufer, 1997; Nation, 2001; Wesche and Paribakht, 2000) due to the recurrent importance given to the focus on forms (FonFs) approach (i.e., form-focused instruction). The rise of this view is deeply rooted in the consideration of the shortcomings of extensive reading, limitations of learners' cognitive processing (VanPatten, 1990) and advantages of noticing the forms (Schmidt, 1990). This implies that form-focused vocabulary instruction can benefit L2 learners because it promotes learner awareness by helping them notice unfamiliar words and thus induces the elaborate

processing of new words. For these reasons, Laufer (2003) questioned the effectiveness of extensive reading and suggested the use of shorter texts complemented with focused vocabulary tasks because it yields better results than reading-only activities. This means that learners need to first encounter unknown words incidentally through the medium of reading (or listening) and then deal with them later by performing word-focused tasks. This approach to vocabulary instruction is also adopted in the present study.

### **2.3.2. L2 Vocabulary Learning through Audio or Audiovisual Input**

Although research with L2 learners regarding their vocabulary gains primarily relied on reading-based approaches, the importance of other modes of exposure to input has been emphasized with more recent studies that benefitted from audio tracks or videos (Brown, Waring and Donkaewbua, 2008; Lin, 2010; Marefat and Hassanzadeh, 2014; van Zeeland and Schmitt, 2013; Vidal, 2011). The reason behind the use of the aural medium to teach L2 vocabulary lies in the hypothesis that reading and listening share the same underlying processes although one includes the comprehension of a written text and the other involves understanding spoken discourse (Gernsbacher, Varner and Faust, 1990; Kintsch and Kozminsky, 1977; Walker, 1976). This notion paved the way for the implementation of listening tasks in L2 classrooms in order to observe if the two input modalities are identical with respect to their contribution to L2 vocabulary acquisition.

A few leading studies have attempted to compare the two mediums of input presentation (i.e., textual and oral). One example is Brown, Waring and Donkaewbua's (2008) study with 35 Japanese pre-intermediate or intermediate EFL learners studying English literature at a university in Japan. They compared vocabulary learning from reading only, reading-while-listening and listening only. Their treatment involved three groups and lasted for 6 weeks. With three graded readers, the learners were exposed to a total of 84 non-English words (28 words from each book) through three input modes. To put it simply, the groups simultaneously covered the same book either by reading it, listening to it or both and the order changed when a new story was assigned. To measure any possible gains in meaning, a translation test (meaning recall) and a multiple-choice test (meaning recognition) were administered. On the immediate post-test, the reading-while-listening (48%) and the

reading-only groups (45%) recognized significantly more of the 28 words learned per book than the listening-only group (29%). On the translation test, the groups respectively supplied a correct meaning for 15%, 16% and 2% of the 28 words learned. This finding implied that recognition occurs earlier than recall and listening leads to fewer gains in word knowledge than reading. However, the scores on the recognition test remained steady after three months, but the scores on the recall test dropped significantly over time, particularly for the reading groups. This finding led to the conclusion that knowledge acquired from listening is more resistant to being forgotten, yet given the limited number of participants and the very low scores of the listening group, it is hard to say that listening accounts for this difference all alone.

To explore whether reading and listening can equally affect the incidental learning and retention of 36 target words among four different proficiency groups, Vidal (2011) conducted a comprehensive study with 248 university freshmen taking a mandatory ESP (English for Specific Purposes) course in Spain. The students were separated into three conditions: one read three academic texts, another listened to three lectures and the other served as the control group who performed neither. After their treatments, the experimental groups also answered some true/false questions and cloze tests for comprehension. The immediate vocabulary post-test scores (a self-reporting vocabulary scale of partial or full word knowledge) showed that reading and listening groups both expanded their lexicon. This progress increased within higher levels of proficiency groups, but reading led to more significant word gains than listening across all levels. However, the one-month-later delayed post-test signalled the opposite. Highly proficient students in the reading condition retained fewer words than highly proficient listeners, while the superiority of reading was still maintained among the lowest language levels. This finding implied that at high proficiency levels words are more vulnerable to decay when acquired through reading and less likely to be forgotten when learned through listening, while the reverse is true for lower-proficiency groups.

A further listening-based vocabulary study was conducted by van Zeeland and Schmitt (2013) to explore the acquisition of form recognition, word class recognition and meaning

recall with four listening passages. Their study included 30 high-intermediate or advanced EFL learners from different L1s enrolled in a postgraduate program at a British university. To prevent any possible testing effects, the learners were divided into two groups: 20 learners took an immediate post-test following the input, while 10 of them took a two-week delayed post-test after receiving the input. Multiple-choice recognition tests were used to measure the knowledge of word form and class. After listening to 24 pseudo words used as target items, the learners indicated their familiarity with the forms and parts of speech of these words. As for the knowledge of meaning, the learners were asked to write an L1 definition or L2 explanation for each target word. Of all 24 target items, the learners recognized 45.8% (11) of the correct word forms and 33.7% (8.1) of the correct grammatical classes on the immediate post-test. However, only 8.5% (2.05) of meaning recall was achieved. The vocabulary gains on the delayed post-test were even lower: 25% (6) for word form, 24.6% (5.9) for word class and 7.5% (1.8) for meaning. These findings showed that the knowledge of word form and grammar was learned much faster than meaning, but much of this knowledge significantly decreased over time. Although the knowledge of meaning was harder to be acquired through listening, it seemed to be retained more over time.

In addition to reading and listening-based studies, previous research also concentrated on the effects of input modality on L2 lexical acquisition by incorporating the audiovisual input. For instance, Lin (2010) wanted to know what roles reading and listening play in incidental vocabulary acquisition along with the use of video. A group of 82 Taiwanese EFL first-year college students participated in the study and formed the three proficiency levels: high reading and high listening (RHLH), high reading and low listening (RHLL) and low reading and high listening (RLLH). Each group watched five video-based lessons, did five gap-filling vocabulary tests and took a word recognition post-test of 15 target words in the end. The gap-filling test scores pointed out that the learners who are highly proficient both in reading and listening significantly outperformed the highly proficient listeners, but there was no significant difference between the RHLH and RHLL groups. Likewise, a comparison of the pre-test and post-test scores also highlighted the significant

progress that the RHLH and RHLL groups achieved in their word recall in contrast to the RLLH group. Therefore, it was suggested that the ability to read better can foster L2 vocabulary learning greater than having good listening skills only.

The influence of input modality on incidental vocabulary learning was further investigated by Marefat and Hassanzadeh (2014). They exposed 63 Iranian English majors to different modes of input (written, aural and audiovisual) by means of three types of materials (reading texts, audio podcasts and vodcasts). Over a period of one month, the learners in Group A watched four news vodcasts; the learners in Group B listened to four news podcasts and the ones in Group C read four political news texts. The learning and retention of 20 target words (5 words per session) were measured through passive recall post-tests (supplying an L1/L2 definition). Each group was presented with an immediate post-test at the end of each session. The delayed post-tests were administered one week after another at the beginning of each subsequent session. Data analysis indicated that the reading and listening groups exhibited no significant difference in their initial gains and retention of the target items, but the vodcast group significantly outscored the others on both immediate and delayed post-tests. Several justifications were made to explain the superiority of the audiovisual input. One reason may arise from the overwhelming demands of the written input on attention mechanisms. When learners are faced with a reading text all at once, they may not notice the unfamiliar words or may simply choose to ignore them for the sake of the gist. While listening, they may also fail to follow the flow of the speech if they cannot recognize the helpful paralinguistic features in the discourse. However, rich contextual cues, authentic language and an engaging content in audiovisual materials can compensate for all these shortcomings that written and spoken texts cannot overcome alone.

As the scarcity of research cited above points out, both spoken and audiovisual input have attracted less attention compared to textual input in incidental vocabulary acquisition. Very few studies have directly addressed the potential impact of listening-only activities on L2 vocabulary learning (Brown, Waring and Donkaewbua, 2008; Vidal, 2011; van Zeeland and Schmitt, 2013). The reason for this lack of research might rely on the finding indicating that learners seem to gain fewer words from listening. However, it was also revealed that

listening has more durable effects on meaning recall. This means that listening might not be the most effective medium for the initial learning of new words, but it can lead to better retention in the long run. Earlier studies also showed that learners find it easy to memorize and recall the target words when presented with the material associating words with sounds; hence catching their attention through both visual and auditory channels (Hu and Deng, 2007; Jones, 2004). Similarly, the audiovisual input has received the same lack of interest in L2 vocabulary studies, but previous research on the use of multimedia materials has concluded that learners obtain significantly higher vocabulary gains when provided with any texts accompanied by videos, pictures or graphics (Akbulut, 2007; Kim and Gilman, 2008; Marefat and Hassanzadeh, 2014). Overall, considering the small number and different contexts of the former studies, this kind of input needs further investigation to determine how it can foster L2 vocabulary knowledge when implemented within word-focused tasks.

## **2.4. Depth of Processing**

### **2.4.1. The Multistore Model**

One pioneering framework concerning the relationship between verbal learning and human memory is the multistore model (also known as the modal model) proposed by Atkinson and Shiffrin (1968). It is a structural model dominated by the idea of three distinct memory stores (a sensory register, the short-term (primary) memory and the long-term (secondary) memory) and the information being transferred from one store to another in a linear way. According to this model, man is perceived like a computer in terms of information processing. The three stores constitute the vital structure (or the hardware) of the system. To process information, stimuli should be detected first by the sensory organs and then registered into the sensory store. If the processor pays attention to the stimuli, the short-term store (STS) can encode and hold the verbal information in an auditory or phonemic format. However, the STS has limited capacity, so the information might be lost within 30 seconds if it is not rehearsed or repeated. The long-term store (LTS), on the other hand, has no limit to its capacity to store information. Once the information is semantically coded in

this type of memory through repetition and rehearsal, it can linger in there from minutes to years.

Although the multistore model provided a valuable insight into the structure of the STS and generated further research into memory, it has been criticized as well. First of all, it regards the STS and the LTS as single units, but Baddeley and Hitch (1974) argued against this unitary structure of the STS with their working memory model. They proposed that the short-term memory consists of several units for different types of information (the central executive, the phonological loop, etc.). Similarly, Tulving (1972) stated that the long-term memory can be divided into different subtypes: episodic (memory for events), semantic (memory for language use) and procedural (memory for actions). In addition, Tulving and Patterson (1968) questioned the consecutive order of the information flow in the multistore model. They opposed the condition that information must necessarily pass through the short-term memory to be able to enter into the long-term memory. Finally, Craik and Lockhart (1972) stated that the model is structurally centred on attention and rehearsal only, so it oversimplifies the processing of the human cognition. Consequently, to address all these problems related to the multistore model, they proposed the depth of processing model.

#### **2.4.2. Levels of Processing**

With their depth of processing framework in the cognitive field of psychology, Craik and Lockhart (1972) aimed to explain how the degree of mental (cognitive) effort exhibited during the course of processing new information can foster its long-term retention. The central idea of the model emphasizes that information is encoded in two levels (shallow or deep) and the degree of decay or retention is determined by the shallowness or depth of cognitive processing. Type I processing happens if the information is processed at the shallow level in the form of visual or auditory traces, but Type II processing can take place only when the information is semantically processed at the deeper level. This means that new information is more likely to be stored in the long-term memory when it is processed deeply during its first encounter. Therefore, when attending, noticing or rehearsing a piece of information, it is crucial to process it at a deeper cognitive stage so that its memory trace

can last longer. Craik and Tulving (1975) further expanded the framework by adding the component of elaboration or richness with which new information is initially encoded. That is how the stimulus is enriched by a wide range of properties such as denotation, connotation and association. If the new information is associated with the existing knowledge by establishing strong links through similarities and contrasts, it will be more accessible later. However, without this elaborate processing, it is highly unlikely that it will be stored in the long-term memory.

Unlike the multistore model, the levels-of-processing framework is a non-structural information-processing model. It views the human memory as the by-product of cognitive processing, so it does not aim to dwell on the distinction between the short-term and long-term memories. Instead of dealing with the structure of memory, it focuses on the underlying mental processes. While shallow processing cannot exceed beyond the structural and phonemic representations of the stimuli, deeper processing can lead to a more meaningful analysis of the information. Therefore, it results in better recall.

#### **2.4.3. Depth of Processing and L2 Vocabulary Acquisition**

Based on the promising nature of the framework, Laufer and Hulstijn (2001) decided to incorporate the notions of depth of processing (Craik and Lockhart, 1972) and elaboration (Craik and Tulving, 1975) into L2 vocabulary acquisition. They thought that if L2 learners elaborate on unfamiliar words (new information) more semantically (at a deeper level), their chance of retrieving these words (retention rate) will increase. One reason for their prediction stems from the idea that learners engage in a deeper process when grasping the meaning of unknown words (semantic processing), but they remain at a shallow level when learning how to write their orthographic forms (structural or phonemic processing). As it is mentioned, semantic processing occurs at a deeper level than structural or phonemic processing, so learners are inclined to recall meaning more than form if they reach the deeper level. The depth of this processing is not determined by the duration with which a new word is held in the short-term memory, but by the degree of the depth or shallowness with which a learner first makes sense of the word. When this analysis gets more meaningful, word recall will be more possible.

The validity of the framework has been questioned by several scholars (Baddeley, 1978; Eysenck, 1978; Hulstijn and Laufer, 2001; Nelson, 1977). One problem with the depth of processing model was found to be its oversimplified view on the processing depth. As Baddeley (1999) implied, learning of new words is not influenced by their orthographic, acoustic or semantic features only, and these three facets of a word cannot be grouped into only three levels of memory trace. The model only takes the cognitive dimension of processing into account, but there are other variables such as motivation, learner strategies, etc. that influence learning. Lockhart and Craik (1978) also acknowledged that their model lacks a clear-cut definition for the notions of depth and elaboration. They stated that although these constructs seem to be independent, circularity is inherent in the model, so it should be regarded as a heuristic framework. Later, Laufer and Hulstijn (2001) also criticized the model because it does not clearly describe what constitutes a processing level and how learners process new words at each level. They noted that the construct of depth requires a precise definition and a concrete measure to be able to determine the superiority of one level over another. They attempted to solve this ambiguity by proposing the Involvement Load Hypothesis, which established the grounds of the present study along with the depth of processing framework.

## **2.5. The Involvement Load Hypothesis**

Based on the levels-of-processing model (Craik and Lockhart, 1972) and elaboration (Craik and Tulving, 1975), Laufer and Hulstijn (2001) intended to formulate a construct to operationalise the notion of processing depth and make it more observable and measurable for L2 vocabulary learning. Their primary goal was to create a task-specific measure for task design that can be used to evaluate the efficacy of particular tasks in terms of lexical retention. As a result, they proposed the Involvement Load Hypothesis (ILH) that helps determine the cognitive load of word-focused tasks by claiming that word retention is conditional upon the degree of cognitive processing and the amount of learner involvement a task induces. If a task is engaging and learners are cognitively involved, there is a greater chance that new words learned through the task will be recalled better.

### 2.5.1. Components of the Involvement Load Hypothesis

The ILH consists of three main components: need, search and evaluation. Need is the motivational and non-cognitive dimension of the hypothesis. It corresponds to the need for achievement. It should be noted here that this need should not be negatively interpreted as the fear of failure. It should be positively viewed as the motivation to fulfil the task requirements for the sake of learning the target words. Based on the source of this need, two degrees of prominence are proposed for the component. Need is hypothesized to be moderate if it is imposed by an external agent such as the teacher. An example is the need to fill in a word in a sentence because the teacher has demanded it. Need is strong when the learner is intrinsically motivated to meet the task requirements (e.g., when the learner makes a decision to look up a word to convey a message when writing a composition).

Search and evaluation are the two remaining aspects of the hypothesis regarding form-meaning associations. Search represents the need to figure out the meaning of an unknown word by looking it up in a dictionary or consulting to external sources such as the teacher or peers. This dimension is distinguished from need and evaluation because it is not labelled under such relative degrees as moderate or strong. Instead, it may be either absent or present. Search is absent when a learner is supplied with the definitions of the target words through such means as marginal glosses or by an authority like the teacher. Similarly, it does not exist when learners can benefit from contextual clues for word-meaning inference during a task. However, search occurs when a learner makes an attempt to discover the meaning of a specific word through dictionary use.

Evaluation entails making the right decision to find the appropriate meaning or correct form of a particular word in relation to the other words in a given context or its further definitions in a dictionary. Moderate evaluation requires distinguishing between the multiple meanings or forms of a lexical item to find the best one that fits the context (as in a homonym, like a swimming pool or pool as a game played with coloured balls). It can also refer to comparing several words with each other and choosing the right option for the specific context (as in matching or gap-filling tasks). Strong evaluation, on the other hand,

involves combining the new word with several others to create original sentences or texts (as in sentence-writing tasks or writing a composition).

By grading vocabulary tasks based on their inclusion of each component, Laufer and Hulstijn (2001) aimed to assess the efficacy of a task in terms of its contribution to the retention of new words in an incidental setting. Based on the absence/presence of the three components and the degree of cognitive processing they require of a learner while performing a task, each component receives a score called an involvement index. This index can be measured through such points or symbols as 0 or (-) (i.e., the component is absent), 1 or (+) (i.e., it is moderately present) and 2 or (++) (i.e., it is strongly present). The involvement index or load of a task is determined by the sum of these scores, so the total index of any task can possibly range from 0 (minimum) to 5 (maximum) (see Table 2.2 for sample tasks). For example, a reading comprehension task with the target words glossed in the margin gets an involvement index of 1 because need is moderate (1) (induced by the text), and search and evaluation are absent (0) (as learners are not asked to find or negotiate the meaning of the target words). However, a sentence-writing task with the glossed words receives an involvement index of 3 because need is moderate (1) (probably induced by the teacher), search is absent (0) (since the definitions of the target words are presented), but evaluation is strong (2) (as learners are required to use the target words to produce novel sentences). With the ILH, Laufer and Hulstijn (2001) claimed that tasks with higher involvement indexes are the most effective ones in yielding greater retention. Based on their claim, the sentence-writing task (3) promotes the learning and retention of the target words more than the reading comprehension task (1) because the former prompts more cognitive effort and learner involvement.

**Table 2.2. Task-Induced Involvement Indexes for Various Tasks**

| Sample Task                                                                           | Need | Search | Evaluation | Index |
|---------------------------------------------------------------------------------------|------|--------|------------|-------|
| 1. A reading comprehension with glossed words but not necessary for the task          | 0    | 0      | 0          | 0     |
| 2. A reading comprehension with glossed words required for the task                   | 1    | 0      | 0          | 1     |
| 3. A reading comprehension with fill-in                                               | 1    | 0      | 1          | 2     |
| 4. A reading comprehension with not glossed words but required for the task           | 1    | 1      | 0/1        | 2/3   |
| 5. Writing sentences with a list of glossed words                                     | 1    | 0      | 2          | 3     |
| 6. Writing a composition with a list of the words (L1 translations are given only.)   | 1    | 1      | 2          | 4     |
| 7. Writing a composition with a list of the words chosen and looked up by the learner | 2    | 1      | 2          | 5     |

*Source.* Laufer and Hulstijn (2001)

### 2.5.2. Empirical Evidence for the Involvement Load Hypothesis

To present empirical support for their hypothesis, Hulstijn and Laufer (2001) conducted two parallel studies with 186 Dutch and Hebrew advanced EFL learners in the Netherlands and Israel. They formed three intact groups in each setting to investigate word retention through three different tasks. Only the evaluation component was differentiated across the tasks, while need and search were held constant. Task 1 involved a reading comprehension with 10 target words or phrases glossed in L1 in the margin. The first group assigned to this task read a text and answered ten multiple-choice comprehension questions. The task induced moderate need (1) because of the relevancy of the glossed items to the questions, but it had no search (0) or evaluation (0) since neither a dictionary look-up nor any comparison of the target items was required. Task 2 (a reading comprehension plus gap-filling) necessitated reading the same passage with the target words omitted and filling them in with the help of the glosses. The addition of gap-filling meant moderate evaluation (1) for the second group. The third group assigned to Task 3 wrote a composition incorporating the target words whose definitions and usages were also presented, so the

task had moderate need (1), no search (0) and strong evaluation (2). Therefore, the three tasks had involvement indexes of 1, 2 and 3 respectively.

On the immediate and delayed post-tests, the learners wrote an L1 equivalent or an L2 explanation for the target items. The results indicated that the composition group significantly outperformed the others on both post-tests in both studies. Although there was no difference between the reading comprehension and gap-filling groups in the Dutch setting, the learners of the fill-in task significantly outsourced those of the glossed reading on both post-tests in the Hebrew experiment. As a result, the study provided strong support for the hypothesis because the composition-writing task loaded with the highest involvement index fostered more vocabulary learning.

Laufer (2003) continued testing the hypothesis through two experiments that involved 142 advanced EFL university students with different L1 backgrounds (Hebrew, Arabic and Russian) in Israel. In these separate studies, a glossed-reading comprehension task with 10 target words (index of 1) was compared to writing isolated sentences or a composition with the words (index of 3). The results of the post-tests in which the learners provided an L1/L2 definition strongly supported the hypothesis. In both measures, the reading comprehension group obtained significantly lower scores than the sentence-writing and composition-writing groups, which verified the efficacy of involving tasks against reading-only tasks. Based on these findings, Laufer (2003) claimed that reading cannot be the major source of L2 vocabulary acquisition because the words learned through reading do not outnumber the words retrieved after completing a word-focused task. Therefore, she suggested that it can be more productive and time-saving to benefit from word-focused tasks to teach vocabulary because in such tasks learners are obliged to notice the unfamiliar words, while such attention can be neglected in longer texts.

Tu (2004) designed three tasks with varying levels of involvement load and probed into their effectiveness on incidental vocabulary learning. The time-on-task factor (i.e., the amount of time invested by a learner to do a task) was also inquired although Laufer and Hulstijn (2001) regarded this factor as an inherent part of the task, thus not open to

manipulation. However, Tu (2004) implied that tasks inducing higher involvement may be more effective not because they trigger a deeper level of processing, but because there is longer exposure to unknown words. With 267 Chinese EFL high school students of diverse proficiency levels, two experiments (time-on-task controlled versus time-on-task uncontrolled) were conducted through a reading comprehension with marginal glosses (Task 1 = 1), a reading comprehension plus gap-filling (Task 2 = 2) and composition writing (Task 3 = 3). The findings showed that when time-on-task was not controlled, word retention was highest in Task 3, lower in Task 2 and lowest in Task 1, which fully justified the hypothesis. When time-on-task was kept constant, the more involving tasks (Task 2 and Task 3) could outrun the least involving task (Task 1), but the difference was not statistically significant. Therefore, when more time was allocated, more words were acquired, so Tu (2004) concluded that the ILH cannot be the mere explanation for task efficacy as time-on-task can also play a facilitative role in vocabulary learning.

With 79 undergraduate beginning learners of Spanish as L2, Keating (2008) replicated Hulstijn and Laufer's (2001) original study to explore whether their predictions can be validated both in active and passive word knowledge for low-proficiency learners. He implemented three tasks similar to the ones devised by Hulstijn and Laufer (2001), but instead of low-frequency items, he preferred eight Spanish pseudowords. The three tasks were a reading comprehension with marginal glosses (Task 1), a reading comprehension plus gap-filling (Task 2) and writing decontextualized sentences with the target words (Task 3), which differed from the composition-writing task in the original study. The tasks had an index of 1, 2 and 3 respectively. The learners translated eight English sentences containing the target words into Spanish on the active recall post-tests, while they supplied L1 correspondents of the items on the passive recall post-tests. Data analysis revealed that partially in line with the hypothesis, Task 2 and Task 3 promoted more passive word knowledge than Task 1 although Task 3 was not most productive. However, fully verifying the hypothesis, the scores of immediate active recall were highest in Task 3 followed by Task 2 and Task 1, but the delayed post-test showed that Task 3 could not maintain its superiority over time as Task 2 received the highest scores. To explain this finding, Keating

(2008) stated that it is no surprise to observe a decrease in knowledge since the learners were not exposed to the target words or did not review them between the testing periods. Moreover, sentence writing may not trigger the kind of deep processing that occurs in composition writing due to not generating a discourse. Nonetheless, Keating (2008) concluded that the hypothesis can be generalized to low-proficiency learners as well.

Following Keating (2008), Kim (2011) carried out two experiments with 104 ESL learners from different L1s to test the hypothesis across various language levels. Based on their TOEFL scores, the learners enrolled in an Intensive English Program (IEP) or an undergraduate program in the U.S. were divided into two groups: less proficient (IEP) and more proficient (Undergraduate). In the first experiment, three tasks (a reading comprehension with graphic organizers (1), a reading comprehension plus fill-in (2) and writing a composition (3)) were compared across levels. The tasks resembled in design to the ones created by Hulstijn and Laufer (2001), but Kim (2011) added graphic organizers to the first task to balance time-on-task. For the second experiment, two tasks (sentence writing and composition writing) inducing the same involvement load (3) were prepared to verify Laufer and Hulstijn's (2001) claim that tasks sharing an identical load of involvement are equally effective in increasing L2 word retention. In these tasks, while one group wrote a composition with the target words, the other wrote original sentences with them.

The data collected through Paribakht and Wesche's (1993) Vocabulary Knowledge Scale (VKS) showed that on the immediate post-test of the first experiment, the composition groups in both levels scored significantly higher than the gap-fill groups which in turn scored higher than the reading groups. On the delayed post-test, the composing groups in both levels significantly outperformed the gap-fill groups which also showed a significant difference from the reading groups. As a result, Kim (2011) noted that strong evaluation may have a far greater influence in the overall load although Hulstijn and Laufer (2001) considered all the components to be identical in value. As for the findings of the second experiment, the composition and sentence-writing groups in both levels obtained very similar scores. Therefore, both experiments were in harmony with the hypothesis since the

more involving tasks led to more retention and the equally-loaded tasks yielded similar word gains. Thus, Kim (2011) implied that lower-proficiency learners can also benefit from involving tasks even though these tasks induce deeper cognitive processing that may challenge such learners' language abilities.

In a study with 122 Taiwanese lower-intermediate EFL learners, Lu (2013) aimed to determine whether gap-filling or composition writing enhances L2 vocabulary learning more. Through four tasks (gap-filling of the target words for once, gap-filling of the target words for three times, a summary plus gap-filling of the target words and writing a summary with the target words), nine words were presented to the learners and their retention was tested by means of active recall, passive recall and sentence production post-tests. On the immediate post-tests, the triple gap-filling task significantly outscored the others, but on the delayed post-tests it indicated no significant difference. As a result, Lu (2013) concluded that multiple gap-filling tasks may not have a long-term effect, but they are more productive than composition tasks for lower-intermediate learners due to time constraints in EFL classrooms.

In most studies associated with the ILH, tasks were exclusively created by manipulating the numerical value of the evaluation component because search can only exist as absent (0) or present (1) but not strong (2). Another factor for this tendency in task design is that many learners are not skilful or knowledgeable when it comes to consulting to a dictionary, which is probably because of the lack of proper training for dictionary use. Therefore, in the preliminary studies search was predominantly kept absent (Hulstijn and Laufer, 2001; Keating, 2008; Kim, 2011; Lu, 2013; Tu, 2004). Apart from these studies that explored the influence of evaluation along with L2 proficiency, there have been several others that investigated the insertion of search through dictionary use across different proficiency levels (Mármol and Sánchez-Lafuente, 2013; Teng, 2015; Yaqubi, Rayati and Gorgi, 2010).

A large amount of data on the ILH has been gathered from advanced learners who are often competent enough to figure out the unknown words with no or little struggle, so Teng

(2015) intended to test the hypothesis across different levels by inserting the search component as well. He also noted that earlier studies largely assessed passive knowledge rather than actual performance. To address these problems, he conducted an experiment with 180 Chinese English majors at 2000, 3000, 5000 or university word level. After doing one of the four tasks (a reading comprehension with glosses (1), a reading comprehension plus blank-filling (2), a reading comprehension with glosses plus writing a composition (3) and a reading comprehension with dictionary look-up plus composition writing (4)), the learners took an immediate and delayed post-test of active word recall and sentence production. The results supported the hypothesis in that the fourth task with the highest involvement load led to the greatest retention, but this was only valid for higher-proficiency groups. This discrepancy implied the emergence of a lexical threshold that allows some learners above a certain level of competence to recognize and comprehend words in authentic materials. As a result, Teng (2015) suggested implementing highly-loaded tasks to more proficient learners because it is challenging for low-level learners to acquire new words incidentally when their reading comprehension is insufficient.

As it has been evident in further studies (Mármol and Sánchez-Lafuente, 2013; Yaqubi, Rayati and Gorgi, 2010), the existence of search has not always confirmed the fundamental prediction of the ILH that tasks with a greater level of involvement enhance better vocabulary learning and retention. For instance, in a study with 69 Iranian intermediate EFL learners, Yaqubi, Rayati and Gorgi (2010) administered three tasks: a reading comprehension with target word look-ups (3), a reading comprehension plus blank-filling (2) and a reading comprehension plus composition writing (3). They found out that the fill-in group significantly outscored the reading-plus-look-up group although the latter performed a task with a higher index. Likewise, the composition task produced significantly better results than the reading group despite their identical level of involvement, which supported the superiority of evaluation over search as Kim (2011) suggested. As a result, the study concluded that an output-oriented task (composition writing) appeared to be more productive than an input-oriented task (reading comprehension plus dictionary use) in spite of their index equivalency, which contradicted

Hulstijn and Laufer's (2001) claim that no task type, whether input or output, is superior to another given that they share the same level of involvement.

Experimenting with search has been proven to be futile by Mármol and Sánchez-Lafuente (2013) in their study with 28 Spanish fifth graders studying English at the elementary level. The learners were assigned to four tasks: Task 1 (a reading comprehension with glosses), Task 2 (a reading comprehension plus gap-filling), Task 3 (writing sentences with glosses) and Task 4 (writing sentences with a dictionary) which differed since it incorporated search through the use of a bilingual dictionary to look up the definitions of 18 target words. The tasks had the indexes of 1, 2, 3 and 4 in order. To test both productive and receptive word knowledge, two post-tests which called for the supply of correct forms and meanings were administered. The scores of the receptive post-test were in favour of Task 3 (sentence writing only) although it did not induce the greatest involvement. While Task 2 and 4 yielded similar gains, Task 1 produced the lowest scores. As for productive word knowledge, Task 3 was significantly most productive again and Task 4 outscored Task 1 and 2. This was explained by task nature in that Task 2 is a receptive task, while Task 4 is a productive one. The type of knowledge the latter task reinforces can account for its superiority on the productive post-test. Therefore, Mármol and Sánchez-Lafuente (2013) warned against the possible effects of task nature. They also concluded that higher degrees of involvement can be too confusing or too demanding for low-proficiency learners due to their cognitive capacity and maturity. Dictionary use also requires certain skills that may be overwhelming for such learners due to their linguistic incapability.

### **2.5.3. The Link between Input Type and Task-Induced Involvement Load**

Most of the research on the ILH has centred on the concern over the influence of varying degrees of search and evaluation on task effectiveness through reading. However, several studies have also probed into the link between the input mode and task-induced involvement load (Fatalaki, 2014; Hassanzadeh, 2016; Jing and Jianbin, 2009; Maleki, 2012). In one of these studies, Jing and Jianbin (2009) investigated whether listening comprehension tasks loaded with different degrees of involvement result in better vocabulary learning. To a group of 86 Chinese non-English major EFL learners of equal

proficiency, they applied three tasks (a listening comprehension with glosses irrelevant to the task (0), a listening comprehension with glosses relevant to the task (1) and a listening comprehension plus composition writing (3)). Two listening passages and 15 target words were selected. The results of the immediate and delayed post-tests in which the learners were supposed to provide a Chinese equivalent or an English definition for the target words indicated that tasks with greater indexes lead to better word gains as the composition-writing group outperformed the others. In a more comprehensive study with 80 Iranian pre-intermediate English majors, Maleki (2012) implemented similar tasks involving 12 listening passages and 40 target words and concluded that tasks with a higher involvement load result in greater retention of the target words.

The effect of input type on task-induced involvement load was also explored by Fatalaki (2014) in another study in which oral and written tasks were compared. To 36 Iranian intermediate EFL learners, he employed three tasks: a multiple-choice task with glosses (1), a fill-in task with glosses (2) and a composition-writing task without glossing (3). The tasks were implemented through the medium of both reading and listening to teach 10 unknown words. The findings revealed that tasks designed with the written input are more productive in the retention of new words than listening-based tasks.

On the other hand, Hassanzadeh (2016) found out that tasks involving the audiovisual input are more effective than reading comprehension tasks. In his study with 21 Iranian undergraduate EFL students, three tasks (Task A = a comprehension task with marginal glosses with an index of 1, Task B = a fill-in task with an index of 2 and Task C = a sentence-writing task with an index of 3) were applied through two input types over six sessions. The written input was provided by means of reading thematically related texts, while the audiovisual input was presented through watching topic-related TV talk shows. The learners completed the tasks through both modes (one task in one mode per session). An active word recall post-test was administered at the end of each session to measure the learning of 10 target words taught per task in two modes (30 words in total). The results indicated that active word recall was significantly influenced by both input modes, but the average of overall word gains from the three tasks was higher in the audiovisual input (4.07

of the 10 words learned) than in the written input (3.54 of the 10 words learned). This finding provided strong evidence for the superiority of using videos over reading texts. When the tasks were compared within each input presentation, the sentence-writing task ( $M = 3.8$ ) was found to be most effective in the written condition followed by the gap-filling task ( $M = 3.4$ ) and the comprehension task ( $M = 3.3$ ). In the audiovisual mode, the sentence-writing task ( $M = 4.4$ ) yielded the greatest word recall again followed by the comprehension task ( $M = 3.9$ ) and the gap-filling task ( $M = 3.7$ ). Although the findings implied that the efficacy of a task is found not to be significantly influenced by input type, Hassanzadeh's (2016) study largely supported the hypothesis that highly involving tasks that rely on the audiovisual input can promote the learning of new words better. The present study also aims to explore this issue by following a similar procedural design implemented across two different language levels.

## 2.6. Research Questions

The preliminary research on the task-induced involvement load has predominantly looked into the effectiveness of tasks with different degrees of search and evaluation (Fatalaki, 2014; Hassanzadeh, 2016; Hulstijn and Laufer, 2001; Jing and Jianbin, 2009; Keating 2008; Maleki, 2012). However, there have been a few studies focusing on tasks with an identical index (Hassanzadeh, 2016; Kim, 2011; Mármol and Sánchez-Lafuente, 2013; Yaqubi, Rayati and Gorgi, 2010). Moreover, previous studies have primarily probed into task efficacy through the written input. Very few studies have added the oral or audiovisual presentation of the input into word-focused tasks (Fatalaki, 2014; Hassanzadeh, 2016; Jing and Jianbin, 2009). Therefore, the present study aimed to investigate the effect of input type (written or audiovisual) on the efficacy of tasks with various or identical degrees of involvement load by measuring both receptive and productive word knowledge. It also looked at the link between input type, task effectiveness and language proficiency. The following research questions guided the study:

- 1) Do vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at lower-intermediate proficiency level

- a) in terms of productive word knowledge?
- b) in terms of receptive word knowledge?
- 2) Do vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at lower-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 3) Do vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at upper-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?
- 4) Do vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at upper-intermediate proficiency level
  - a) in terms of productive word knowledge?
  - b) in terms of receptive word knowledge?

## **CHAPTER III: METHODOLOGY**

The present chapter gives detailed information concerning the design and methodology of the present study. It explains the background and characteristics of the participants in addition to the materials, vocabulary tasks and testing instruments that were utilized during the course of the study. It concludes by describing the scoring protocols and statistical procedures that were employed for data analysis.

### **3.1. Research Design**

The present study was conducted in the spring term of 2016-2017 academic year. Based on the predictions of the Involvement Load Hypothesis (Laufer and Hulstijn, 2001), it aimed to explore the relationship between task-induced involvement load and input type and their mutual effects on incidental vocabulary learning across two separate proficiency levels. The study adopted a quasi-experimental design with eight experimental and four control groups through the implementation of six vocabulary tasks during one session and administration of vocabulary pre- and post-tests at the beginning, immediately after the treatment and two weeks later.

### **3.2. Participants**

The participants consisted of 236 (116 male and 120 female) Turkish undergraduate EFL learners enrolled in a mandatory Intensive English Program at a preparatory school of a state university in Turkey. They were young adults ranging from 18 to 22 in age. They had been attending English classes at the institution for 5 months to be proficient enough for their departmental courses that would be delivered in English in the following year. The participants' majors included Engineering, Business Administration and International Relations. The language program that the participants were enrolled in was based on a modular system which required them to meet the particular criteria in order to pass through four eight-week modules (from A1 to B2 defined by the Common European Framework of Reference for Languages) to complete the program. The common requirements for each module involved completing a number of written and oral assignments for a portfolio and taking pop quizzes, writing tests, midterms and final examinations.

The study involved twelve intact classes to which the participants from two different levels of proficiency were randomly assigned by the administration. Convenience sampling was used while selecting these classes to form the experimental and control groups. Six classes ( $n = 123$ ) were at the lower-intermediate level, while the other six ( $n = 113$ ) were upper-intermediate. The participants were considered to be at these levels of proficiency based on their scores on a 40-item multiple-choice placement test by Oxford University Press that was administered earlier that year and the institutional examinations that were carried out during the fall semester. From two different language levels, 75 participants were assigned to the sentence-writing tasks (by reading or watching), 78 participants were assigned to the gap-filling tasks (by reading or watching) and 83 participants were either in the reading or watching control groups (see Table 3.1 for the distribution of participants). All of them volunteered to participate in the present study, but they were not informed about its purpose or the upcoming vocabulary post-tests not to encourage any memorization of the target words.

**Table 3.1. Distribution of Participants**

| <b>Task Type</b>             | <b>Lower-Intermediate</b> | <b>Upper-Intermediate</b> |
|------------------------------|---------------------------|---------------------------|
| Reading only                 | 21                        | 19                        |
| Watching only                | 23                        | 20                        |
| Fill-in by reading           | 20                        | 17                        |
| Fill-in by watching          | 20                        | 21                        |
| Sentence writing by reading  | 21                        | 19                        |
| Sentence writing by watching | 18                        | 17                        |
| Total                        | 123                       | 113                       |

### **3.3. Data Collection Instruments**

Several instruments were used in the study for research purposes. First of all, two different video podcasts as audiovisual materials and two separate reading passages as written texts were chosen for the study. A total of 16 vocabulary items were chosen as target words to be used in the vocabulary tasks. Finally, vocabulary pre-tests and post-tests were implemented to gather data for the present study.

### 3.3.1. Video Podcasts

Two different video podcasts were selected as audiovisual materials for the present study. Both of them were chosen from a well-known skills-integrated language coursebook series in Turkey, namely *New Headway (4<sup>th</sup> edition) Series*, which were being covered by the learners at the time of the study. The videos accompanied by their printable activity worksheets were part of a built-in section of the projectable digital supplementary resources of the books. The video assigned to the lower-intermediate learners was about a county in England named Cornwall and a volunteer organization that is situated there. The video lasted around 10 minutes and consisted of 1170 words. The other video that the upper-intermediate learners watched was related to the city of Vancouver and its action plan to be the greenest city in the world. The duration of this video was 8 minutes. A total of 1055 words were included in the video.

The criteria for choosing these two video podcasts from a set of twenty-four ones in the two coursebooks were based on four reasons. First of all, each coursebook had twelve videos prepared for each unit, but the language used in the selected videos was much easier to follow and understand than that of the remaining videos. Second, the chosen videos provided a sufficient number of target words that were assumed to be unknown to the learners, but the rest of them had either too many frequent words which could not possibly be selected as target words or too many unknown words which would interfere with text comprehension. Furthermore, the two video podcasts were thematically related. Both introduced a county/city by describing a social issue regarding it (volunteer or environmental work). Finally, the researcher chose these two videos because the learners were familiar with their topics as a result of studying the corresponding units in their regular classroom hours. If used, videos from other resources would have caused topic unfamiliarity or language difficulty for the learners.

### 3.3.2. Reading Passages

Two different reading passages for two groups of proficiency were used in the study. Both passages were the transcribed versions of the video podcasts, but they were adapted and a little shortened to look like an academic reading text by excluding some colloquial phrases,

altering several sentence structures and removing some redundant words. This was also done for time management since including every word from the videos would make the texts unnecessarily longer. The reading passage assigned to the lower-intermediate learners was called ‘The Working Week’. It was an expository text about a volunteer organization that is located in Cornwall, England. It contained seven paragraphs with a total of 792 words. The upper-intermediate learners also received a similar reading passage called ‘The Best City in the World’. It was also an expository text about Vancouver and its city plan to become the greenest city in the world in 2020. It consisted of eight paragraphs with a total of 600 words.

### 3.3.3. Target Words

16 target words (8 per level) were selected in total for investigation. Before choosing these target items, all the words that were expected to be unfamiliar to the learners at both levels (12 for lower-intermediate and 15 for upper-intermediate) were selected by the researcher from the two videos and also checked on their appropriateness by two other instructors teaching at both levels. The reason for choosing a greater number of possible target words was to test all of them on their familiarity on the vocabulary pre-tests and select the words that were all unknown to the relevant group of participants. Based on the results of the vocabulary pre-tests, the target words ultimately assigned to the lower-intermediate group included three nouns, three adjectives and two verbs: *pager*, *current*, *pupil*, *dedicated*, *stunning*, *treacherous*, *launch* and *patrol*. As for the upper-intermediate learners, the eight target words consisted of the following nouns, adjectives and verbs: *display*, *creek*, *abundance*, *complacent*, *ultimate*, *conscientious*, *boast* and *implement*. The latter set of words resembled to the former one in their parts of speech to eliminate any possible effects of word class on the learnability of the target items.

### 3.3.4. Vocabulary Tasks

Six classes from each level of proficiency were separately assigned to one of the six vocabulary tasks in the study: a reading-only (RO) task, a watching-only (WO) task, a fill-in-by-reading (FR) task, a fill-in-by-watching (FW) task, a sentence-writing-by-reading

(SR) task and a sentence-writing-by-watching (SW) task. Table 3.2 presents an overview of the tasks and their involvement indexes.

**Table 3.2. Overview of Tasks and Their Indexes**

|             | Control Group Tasks |          | Fill-in Tasks |          | Sentence-writing Tasks |          |
|-------------|---------------------|----------|---------------|----------|------------------------|----------|
|             | Reading             | Watching | Reading       | Watching | Reading                | Watching |
| Need        | 0                   | 0        | 1             | 1        | 1                      | 1        |
| Search      | 0                   | 0        | 0             | 0        | 0                      | 0        |
| Evaluation  | 0                   | 0        | 1             | 1        | 2                      | 2        |
| Total Index | 0                   | 0        | 2             | 2        | 3                      | 3        |

The involvement index of the reading-only and watching-only tasks is 0 because in these tasks need (0) is absent (as the learners do not need to know the target words to be able to answer the comprehension questions that followed the input), search (0) is also absent (as the learners were not asked to find out the definitions of the target words) and evaluation (0) is absent too (since the learners were not required to check the grammatical or semantic appropriateness of the target words in any context). Therefore, these two tasks that were devised without any enhancement techniques or noticing effects were utilized in the control groups to test their productiveness in incidental vocabulary learning compared to the other tasks in the study.

The two fill-in tasks (by reading or watching) shared the same task-induced involvement load with an index of 2 because need (1) is moderate (as the learners needed to refer to the target words to do the task properly), search (0) is absent (as the target words were glossed) and evaluation (1) is moderate (as the learners were asked to complete the blanks by selecting the correct answers from a set of target words). Likewise, the sentence-writing tasks through reading or watching were designed to comprise an identical index (3) with moderate need (1) induced by the task itself, no search (0) with the glossed target words and strong evaluation (2) because the learners were expected to combine the target words with others in a context to create original sentences. The identical involvement indexes within each task type would enable the researcher to see whether input modality had any

effects on the learning and retention of the target words, while the varying levels of involvement load across the three task types would help the researcher see whether tasks with a higher degree of evaluation had any effects on incidental vocabulary learning and retention.

#### **3.3.4.1. Reading-only (RO) Tasks**

After a short in-class discussion designed as a warm-up activity with a general question about the topic of the relevant reading passage, the learners assigned to the reading-only tasks in both proficiency groups were instructed to read the given texts and complete the subsequent comprehension questions. These questions consisted of a wide variety of comprehension-based exercises such as answering open-ended questions, matching numbers with the information they refer to, matching people with their actions, underlining the correct information and completing the sentences with the relevant information. The original supplementary video worksheets of the language coursebooks provided the base for these questions in the reading-only tasks. Upon finishing these questions, the learners were required to discuss another thematically related question with a partner and briefly report their ideas to their peers in the classroom. Overall, it took around 35 minutes for the participants to perform the entire task (see Appendix A for the task layouts). As mentioned before, this task induced an index of 0 because the target words were not bolded, highlighted or underlined not to draw any attention to them. They were not relevant to the comprehension questions either, which means that knowing these words was not required to find the correct answers to the questions. In other words, the learners were not directed to deal with the target items in any way that would help them notice and thus learn these words, so the reading-only tasks induced no learner involvement in terms of the learning and retention of the target vocabulary. As for in-text frequency, the learners were exposed to the target words once during their task, but the words were not highlighted as mentioned.

#### **3.3.4.2. Watching-only (WO) Tasks**

Similar to the reading-only tasks, the watching-only tasks involved the same comprehension questions, but this time the input was displayed through the audiovisual mode. The participants assigned to this task were not exposed to any written texts. Instead,

they watched a video podcast that presented the same content as that of the corresponding reading passage selected for each proficiency level. The subtitles in the video were turned off so that the learners would only have access to the visual or aural cues provided in the video. Otherwise, they would follow the subtitles as opposed to watching the video, which would alter the nature and purpose of the task. Before the main task, the learners were asked to do the same warm-up activity as the reading-only groups performed by discussing a topic-related question. Then they were given 10 minutes to look at the comprehension questions before the first viewing. Each video chosen to be presented to the two proficiency groups was played twice in order to facilitate learner comprehension. After the second viewing, the answers to the comprehension questions were checked in class and a final discussion was carried out to wrap up the task. Overall, the participants were given approximately 40 minutes to finish the whole task. The total involvement index of the task was calculated as 0 since it induced no need (0), search (0) or evaluation (0) (see Appendix B for the task layouts). During the task, the exposure to the target words was only once through auditory channels, and the learners did not need to acquire the meaning of the target items to be able to perform their task.

### **3.3.4.3. Fill-in-by-reading (FR) Tasks**

Fill-in-by-reading groups were provided with the same reading passages and comprehension questions, but they also received a glossary of the target words along with their Turkish equivalents and English definitions because their texts were gapped. After a brief discussion that served as the warm-up activity to make the learners familiar with the topic, the learners in this group were supposed to fill in the blanks in their reading passages with the given target words. They were then asked to share their answers with their instructor and peers to see whether the target items had been placed correctly in their allocated spaces. The main task was followed by the completion of the same comprehension and discussion questions as those of the former groups. The entire task lasted for around 40 minutes and had an involvement index of 2 because need (1) was moderate, search (0) was absent and evaluation (1) was moderate. In other words, the participants in this group were expected to tackle with the target words and compare them

with others to check their appropriateness, but they did not have to look up or search for any targeted meanings (see Appendix C for the task layouts). In terms of exposure frequency, the target words appeared only once in the glossed format during the task.

#### **3.3.4.4. Fill-in-by-watching (FW) Tasks**

Following the warm-up discussion, the fill-in-by-watching groups in both proficiency levels were instructed to look at the provided glossary of the target words for a few minutes. Then they had to fill in these words into the blanks left out in the sentences directly taken from the relevant audiovisual resource while they were watching it. The rationale for extracting intact sentences from the videos was that any other variables such as listening comprehension or recognizing synonyms/antonyms would not interfere with the learners' ability to recognize the target words and write them down into the gapped sentences. After completing the sentences with the target items upon their first encounter with the video, the learners checked their answers in class. The second viewing was carried out to allow the learners to answer the given comprehension questions. Followed by the comprehension check of the material, the learners performed the final discussion and finished the task completely. The entire task ended in about 45 minutes. The involvement index of the task was regarded to be 2 as need (1) and evaluation (1) were moderate, but there was no search (0) because the target definitions were explicitly presented, so the participants only needed to find out the correct sentence each target word fits in (see Appendix D for the task layouts). In terms of frequency, the learners were exposed to the target words twice (with a glossary list and through the video podcast) while completing their task.

#### **3.3.4.5. Sentence-writing-by-reading (SR) Tasks**

In both proficiency groups, the sentence-writing-by-reading tasks were made up of two parts. The first part of the task was quite similar to the reading-only task because the learners were expected to do the small warm-up activity, read the given passage, answer the following comprehension questions and check their answers with the class. This part was added to the sentence-writing task to provide a meaningful context for the learners to realize how the target items were both grammatically and semantically used in original

sentences in the reading passage. The inclusion of the mentioned part would also allow the researcher to balance the variable of time-on-task across various task types implemented in the study. However, there were several differences that distinguished this part from that of the reading-only task. First of all, in the sentence-writing task the target words appeared in a glossary list with their L1 equivalents and L2 definitions and were highlighted in bold print in the passage so that the learners could notice these words and refer to their Turkish counterparts/English explanations while reading their texts. In contrast, the reading-only task did not include any highlighted or glossed items not to draw any particular attention to the target words. The second difference was regarding the number of the questions aimed at checking text comprehension. The comprehension questions of the sentence-writing-by-reading tasks were fewer than the reading-only tasks in order to control the time that the learners spent to perform the former tasks. If the number of questions had not been limited, the learners would have lingered on this part of the task more than needed, thereby having less time left for the most significant part of the task (sentence writing).

The second part of the sentence-writing-by-reading tasks involved writing novel sentences with the target words supplied with their Turkish equivalents and English definitions. On their own, the learners were instructed to form eight original sentences that would somehow reveal the meaning of the eight target words. The part of speech of each word was also written next to the word so that the learners could use the target words grammatically as well. When all the learners completed writing their sentences, several of them were told to read their example sentences out loud so that the meaning of each target word would be covered at least once in class. Following this activity, the whole task ended with the final discussion. On average, it took 50 minutes for the learners to finish the task. The number of exposures to the target words during the task was counted as three (glossed with a list twice and highlighted in the text once). Overall, the sentence-writing-by-reading tasks induced an index of 3 because need (1) was moderate, search (0) was absent, but this time evaluation (2) was strong (see Appendix E for the task layouts). The participants assigned to these tasks had to compare the target words with other words in the target language based on their grammaticality and semantic features in a new context. This would

require a stronger involvement by the learners as opposed to only comparing the target words across one another in the fill-in tasks.

### **3.3.4.6. Sentence-writing-by-watching (SW) Tasks**

The sentence-writing-by-watching tasks in the two proficiency levels also consisted of two sections. The first section resembled that of the sentence-writing-by-reading task in that the learners started with the warm-up discussion before they were instructed to answer several comprehension-based questions. However, in this task the participants did not receive any written texts. Instead, they were presented with the same glossary of the target words. Then they were asked to go over the given L1/L2 definitions before watching the video that corresponded to the relevant reading passage. This part of the sentence-writing-by-watching task finished after the learners watched the video twice, completed their questions and checked their answers in class. The second part of this task required the learners to create eight original sentences that incorporated each of the target words by indicating their meanings to some extent. The whole task ended in about 55 minutes when some of these sentences were read out loud and the final discussion of the theme was held in class. Since the sentence-writing-by-watching tasks induced moderate need (1), no search (0) and strong evaluation (2), the involvement index of these tasks was calculated as 3 (see Appendix F for the task layouts). Finally, the target words appeared three times (glossed twice and heard once) during the task.

### **3.3.5. Vocabulary Tests**

Two weeks before the implementation of the vocabulary tasks, all the participants were instructed to take a vocabulary pre-test prepared exclusively for each level of language proficiency in order to check whether the participants knew the meaning of the target words prior to the study. Each test involved a set of 20 words including the eight target words particularly selected for that level and 12 distracters. Similar to the vocabulary tests used in Hulstijn and Laufer's (2001) study, in these pre-tests the participants were asked to provide Turkish equivalents or English definitions of the words in the list that were written in an alphabetical order. They were also told to put an "X" next to the words that they were not

familiar with. Each correct answer received a full point (1), while an incorrect answer got a score of zero (0). If the answer was partially correct, half a point was granted (see Appendix G for the vocabulary pre-tests).

Two types of vocabulary post-tests (productive and receptive) were administered to all the participants immediately after the completion of the vocabulary tasks and two weeks later. The productive post-tests for each proficiency level consisted of the definitions of the eight target words delivered in both languages (Turkish and English) in an alphabetical order and required the participants to recall and write the target words next to their correct explanations (see Appendix H for the productive post-tests). The tests shared the same definitions as the ones provided in the glossary lists of the vocabulary tasks so that the participants would not have to undergo any other cognitive processes other than retrieving the target items. They were also instructed not to employ any guessing strategies such as supplying the same target word more than once. Instead, they were told to leave out the words that they could not remember at all. However, they were also asked to write their partial answers even if their answers contained nothing more than a few letters or only initials. The rationale behind this instruction was that the present study aimed to measure both partial and full vocabulary gains regarding the knowledge of the word form. On the delayed productive post-tests, the order of the definitions was randomly changed to prevent the participants from taking advantage of the previous order on the immediate post-tests and hence guessing the target forms.

The scores of the productive post-tests were calculated using a lexical production scoring protocol proposed by Barcroft (2000). According to this analytical scoring tool (see Appendix J for the scoring table), points were given based on the complete and partial word knowledge by counting the number of letters correct or the number of letters present and calculating their percentages. A correct letter referred to any letter of a word written in its correct place, while a present letter referred to any letter of a word written but not in its correct place. For example, for the target word ‘current’, the following answers would get a score of .25: ‘c...’ or ‘...t’ (since at least one letter (1 of 7 = 14%) was in its correct place) and ‘urad’ (since at least 2 letters (2 of 7 = 28%) were present but not in their correct

places). A full point (1) would be given to the correct form ‘current’ written next to its correct definition. Based on these criteria, each correct full form of the target words was calculated as 1, while the words not written or written next to an incorrect definition received a score of zero. If a learner wrote the same target word more than once (the word being correct for only one definition), both answers were regarded as incorrect and received a score of 0 because the learner applied a guessing strategy. The maximum score that could be reached on the productive post-tests was 8 points provided that all the target forms were placed next to their correct definitions.

The receptive post-tests included the eight target words presented in an alphabetical order and expected the participants to write the correct L1 equivalents or L2 definitions next to the words (see Appendix I for the receptive post-tests). The participants were asked to provide clear explanations illustrating the meaning of the target words. The aim of the receptive post-tests was to see whether the participants could recall the meaning of the target words. Similar to the delayed productive post-tests, the delayed receptive post-tests consisted of the target words arranged in a different order to prevent the participants from remembering the prior order on the immediate post-tests. The scores on the receptive post-tests were calculated by counting the correct explanations of the target words. A full point (1) was given to a correct explanation written in L1 or L2, while half a point (0.5) was given to a semantically acceptable explanation of the target word (e.g., explaining ‘abundance’ as greatness instead of richness). A score of zero was given to the explanations not written or written next to an incorrect target word. If a participant wrote the same explanation more than once (the explanation being correct for only one target word), both answers were accepted as incorrect because the participant probably used a guessing strategy.

### **3.4. Data Collection Procedure**

Before the implementation of the vocabulary tasks or tests of the present study, the research purpose and procedure were explained to the institution in advance. The instructors who took part in the study were also informed about the aim of the study and their consent was taken. However, the learners were not told in advance that they were participating in a

research study. The reason why these learners were not provided with any information regarding the purpose of the study or the upcoming vocabulary post-tests was not to encourage them to commit the target words into their memory.

Two weeks before the task implementation, the vocabulary pre-test was administered to the participants to test the possible familiarity of the learners with the target vocabulary items. Based on the results of the vocabulary pre-test, the target words were selected from the ones that were unknown to the participants. During Week 2, the researcher informed the other five instructors who were teaching in the target classrooms about how to implement the vocabulary tasks and tests assigned specifically to their classrooms. These instructors were randomly selected because the institution placed each one of them randomly in the classrooms that constituted the research context for the study. The six participant instructors, including the researcher herself, all received enough copies of the vocabulary tasks and tests in paper for their classrooms beforehand and one copy of the answer key to their relevant vocabulary task.

At the beginning of Week 3, all the participants simultaneously completed one of the six vocabulary tasks particularly assigned to their class in their regular classroom hours either by receiving a reading passage or watching a video. Each participant completed his or her task individually in the presence of the instructor and then all the printed materials related to the task (e.g., glossary lists, reading passages, etc.) were collected from the participants and not returned to them. Immediately after the completion of the assigned task, the participants unexpectedly took the productive and receptive posttests in that order. If the receptive post-test had been administered first, the participants could have seen how the target forms were written, so it might have been to their advantage to remember how to write these words. To prevent this recency effect, they took the productive post-test first. After a two-week period (at the end of Week 5), the unannounced delayed productive and receptive post-tests were administered to measure word retention. The implementation procedure remained the same for both proficiency levels.

### 3.5. Data Analysis Procedure

For the purposes of the present study, the data obtained from the aforementioned vocabulary pre-tests and post-tests were analyzed by means of quantitative methods. The scores of the participants on five different kinds of vocabulary tests were compared by using the Statistical Package for the Social Sciences (SPSS) 23.0 within each level of language proficiency. To determine whether parametric tests were appropriate for the quantitative analysis, all variables were tested for the normality of their distribution by using the Shapiro-Wilk test. The significance level for this normality test was set at  $p = .05$ . The null hypothesis for this test is that the data are normally distributed. If the significance value of the Shapiro-Wilk test is greater than .05, the null hypothesis is accepted and the data are assumed to be normally distributed. If it is less than .05, the null hypothesis is rejected and the data significantly deviate from a normal distribution.

Since the data appeared not to be normally distributed based on the result of the Shapiro-Wilk test, the non-parametric Kruskal-Wallis test was used to determine if there were statistically significant differences between the task groups in their immediate and delayed productive and receptive vocabulary post-test scores. This test is considered to be the non-parametric alternative to one-way ANOVA (analysis of variance) for a statistical difference between two or more groups. It is used when the assumptions of one-way ANOVA (the dependent variable is normally distributed and there is approximately equal variance on the scores across groups) are not met. The Kruskal-Wallis test aims to show if there is a statistically significant difference among the medians of three or more data sets rather than taking the mean differences into account as applied in one-way ANOVA. The null hypothesis for this test states that the centre of the distribution for all the data sets is equal to each other. The alternative hypothesis assumes that at least one of the sets has a different measure of centre than the others. Therefore, this test compares the median scores (with their minimum and maximum values) of the relevant data sets to see whether the medians on a dependent variable are the same across all levels of a factor or not.

The Wilcoxon signed-ranked tests were carried out to determine whether there were statistically significant differences between the median scores of the first (immediate) and

second (delayed) implementation of the productive and receptive vocabulary post-tests within each task group. This test is regarded to be the non-parametric equivalent of the dependent sample t-test (or paired-samples t-test). The paired-samples t-test aims to indicate whether the mean difference of two repeated observations is zero and it can be used with metric (interval or ratio) and normally distributed data. When these assumptions are not met, the Wilcoxon test is used. The basic requirements to apply this statistical method are that the data are from the same participants and are paired, they are measured on at least one interval scale, and the paired measures are randomly and independently drawn. If the  $p$ -value of the Wilcoxon test is smaller than .05, then the null hypothesis that the median difference between the two paired data sets is equal to zero is rejected and the median of the differences is significantly different. This also means that the two measurements do not share the same distribution with the same median. However, if the  $p$ -value of the test is greater than .05, then the null hypothesis is accepted and the medians of the two testing occasions are considered to be equal.

## **CHAPTER IV: RESULTS**

This chapter presents the findings of the research implementation. The productive and receptive vocabulary gain results of the lower-intermediate participants are presented in section 4.2, while the productive and receptive vocabulary gain results of the upper-intermediate participants are provided in section 4.3. Finally, the chapter concludes with a summary of the major findings.

### **4.1. Test of Normality**

The lower-intermediate and upper-intermediate participants' vocabulary post-test scores were subjected to the Shapiro-Wilk test of normality to determine whether the implementation of a parametric test was possible. The results showed that both groups did not meet the criteria for a normal distribution due to having a  $p$ -value less than 0.05 (for the results of the normality tests, see Appendix K). Therefore, it was concluded that the data collected from the two proficiency groups were not normally distributed and thus non-parametric statistical procedures were conducted to analyze and interpret the data concerning both groups.

### **4.2. Results for Lower-Intermediate Proficiency Level**

The results belonging to the lower-intermediate participants were provided in the subsequent divisions of this chapter by referring back to the relevant research questions.

#### **4.2.1. Vocabulary Gain from Tasks with Low vs. High Involvement Load**

The first research question aimed to investigate whether vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at lower-intermediate proficiency level in terms of productive and receptive word knowledge. In order to answer this question, the immediate and delayed productive and receptive vocabulary post-test scores of the lower-intermediate learners who performed six variously loaded tasks were compared by means of Kruskal Wallis tests. The results of these tests were presented in the following sections.

#### 4.2.1.1. Gain in Productive Word Knowledge

The median scores of the lower-intermediate learners on the immediate and delayed productive vocabulary post-tests were presented in Table 4.1.

**Table 4.1. Productive Word Gain and Retention at Lower-Intermediate Level (Index Based)**

| Groups | Index | N  | Immediate   |             | Delayed     |             |
|--------|-------|----|-------------|-------------|-------------|-------------|
|        |       |    | $\tilde{x}$ | Min. – Max. | $\tilde{X}$ | Min. – Max. |
| SR     | 3     | 21 | 5.5         | 1.25 – 7.5  | 2           | 0 – 7.5     |
| SW     | 3     | 18 | 4           | 1 – 7.75    | 3.38        | 0 – 8       |
| FR     | 2     | 20 | 5.38        | 1 – 8       | 3           | 1 – 7.75    |
| FW     | 2     | 20 | 2           | 0 – 7.75    | 0           | 0 – 4.5     |
| RO     | 0     | 21 | 0.75        | 0 – 1       | 0.25        | 0 – 3       |
| WO     | 0     | 23 | 0           | 0 – 3.75    | 1           | 0 – 3       |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only. Scores were given out of 8 points.

In order to find out whether the learners who were exposed to the eight target words through six different tasks with low or high involvement loads differed in terms of their immediate productive word gain, a Kruskal Wallis test was conducted to compare their scores on the immediate productive vocabulary post-test. The findings were shown in Table 4.2.

**Table 4.2. Initial Gain in Productive Word Knowledge at Lower-Intermediate Level**

|                   | Groups | N  | $\tilde{x}$ | $X^2$  | p        |
|-------------------|--------|----|-------------|--------|----------|
| <b>Immediate</b>  | SR     | 21 | 5.5         | 76.649 | < 0.001* |
| <b>Productive</b> | SW     | 18 | 4           |        |          |
| <b>Post-test</b>  | FR     | 20 | 5.38        |        |          |
|                   | FW     | 20 | 2           |        |          |
|                   | RO     | 21 | 0.75        |        |          |
|                   | WO     | 23 | 0           |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

Table 4.2 demonstrates that there was a statistically significant difference among the groups in their immediate productive post-test scores ( $p < 0.001$ ). It was observed that the two control groups (reading only and watching only) significantly obtained lower scores than

the other task groups, but no significant differences were found among the experimental groups. With respect to the order of the medians, the SR group had the highest score ( $Mdn = 5.5$ ) followed by the FR group ( $Mdn = 5.38$ ). While the SW group got a median score of 4, the FW group obtained a score of 2. Finally, the RO ( $Mdn = 0.75$ ) and WO ( $Mdn = 0$ ) groups obtained the lowest scores.

As the aforementioned results pointed out, the comparison of the experimental groups based on the involvement load indexes of their tasks showed that the sentence writing groups outscored the gap filling groups when they completed their tasks through the same kind of input mode (i.e., written or audiovisual). For example, the SR group ( $Mdn = 5.5$ ) slightly performed better than the FR group ( $Mdn = 5.38$ ). Similarly, the SW group ( $Mdn = 4$ ) also outperformed the FW group ( $Mdn = 2$ ). This shows that sentence writing tasks (index = 3) lead to greater initial knowledge of the word form than fill-in tasks (index = 2). Both types of tasks also enhance productive word knowledge more than the RO and WO tasks with no involvement load. Therefore, it can be concluded that highly involving tasks are more effective in vocabulary enhancement than tasks with a lower or no degree of involvement load.

#### 4.2.1.2. Retention in Productive Word Knowledge

To see whether the lower-intermediate participants were able to retrieve the eight target forms after a two-week interval, a Kruskal Wallis test was carried out by comparing the learners' scores on the delayed productive vocabulary post-test. Table 4.3 presents the delayed productive vocabulary post-test scores of the lower-intermediate participants.

**Table 4.3. Retention in Productive Word Knowledge at Lower-Intermediate Level**

|                   | Groups | <i>N</i> | $\tilde{x}$ | $X^2$  | <i>p</i> |
|-------------------|--------|----------|-------------|--------|----------|
| <b>Delayed</b>    | SR     | 21       | 2           | 55.726 | < 0.001* |
|                   | SW     | 18       | 3.38        |        |          |
| <b>Productive</b> | FR     | 20       | 3           |        |          |
|                   | FW     | 20       | 0           |        |          |
|                   | RO     | 21       | 0.25        |        |          |
|                   | WO     | 23       | 1           |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0),  $*p < .05$ . Scores were given out of 8 points.

The results indicated a statistically significant difference among the groups ( $p < 0.001$ ). On the second measurement of their productive word knowledge, the SW group ( $Mdn = 3.38$ ) significantly outscored the FW group ( $Mdn = 0$ ) and the control groups as well as performing slightly better than the FR group ( $Mdn = 3$ ). While the SW and FR groups' vocabulary gain demonstrated significant differences with the FW and control groups, the SR group ( $Mdn = 2$ ) also significantly outperformed the RO ( $Mdn = 0.25$ ) and FW groups. No other significant differences were found across the remaining groups. Surprisingly, the FW group achieved the lowest score despite the index of their task.

With respect to the productive word retention of the sentence writing groups who completed the tasks with the highest involvement load (index = 3), it can be suggested that the learners in these groups were still able to recall more of the correct word forms than the others. For instance, the SW group did significantly better than the FW group and scored slightly higher than the FR group. The only exceptions were the FR group (index = 2) who performed slightly better than the SR group (index = 3) and the control groups (index = 0) who outscored the FW group (index = 2) despite having a lower level of involvement load.

#### **4.2.1.3. Changes in Productive Word Knowledge within Groups**

A Wilcoxon signed-ranked test was carried out by comparing the immediate and delayed productive vocabulary post-test scores of the lower-intermediate learners to determine whether these learners were able to retain their productive word knowledge over time (for the results of the test, see Appendix L). The results showed that all the groups experienced a change in their productive word knowledge within two weeks' time, but only four of them were significant. There was a significant difference between the immediate and delayed post-test scores of the SR group ( $Z = -3.304$ ,  $p = 0.001$ ). The learners in this group significantly lost their knowledge of the correct forms. Similarly, the scores of the SW group on the immediate post-test significantly dropped on the delayed post-test ( $Z = -2.419$ ,  $p = 0.016$ ). The FW group also experienced a significant loss in their productive word knowledge ( $Z = -3.632$ ,  $p < 0.001$ ). Interestingly, the WO group significantly increased

their productive word knowledge on the delayed post-test despite not dealing with the target forms at all ( $Z = -3.240$ ,  $p = 0.001$ ).

#### 4.2.1.4. Gain in Receptive Word Knowledge

The median scores of the lower-intermediate learners on the immediate and delayed receptive vocabulary post-tests were shown in Table 4.4.

**Table 4.4. Receptive Word Gain and Retention at Lower-Intermediate Level (Index Based)**

| Groups | Index | N  | Immediate |             | Delayed   |             |
|--------|-------|----|-----------|-------------|-----------|-------------|
|        |       |    | $\bar{x}$ | Min. – Max. | $\bar{X}$ | Min. – Max. |
| SR     | 3     | 21 | 7         | 3 – 8       | 7         | 4 – 8       |
| SW     | 3     | 18 | 7         | 3 – 8       | 7.5       | 2 – 8       |
| FR     | 2     | 20 | 6.25      | 2 – 8       | 6.25      | 0 – 8       |
| FW     | 2     | 20 | 3.75      | 0 – 8       | 2         | 0 – 8       |
| RO     | 0     | 21 | 2         | 0 – 6       | 2         | 0 – 6       |
| WO     | 0     | 23 | 4         | 0 – 6       | 3         | 1 – 6       |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only. Scores were given out of 8 points.

To determine whether the lower-intermediate participants acquired any of the eight target words, a Kruskal Wallis test was conducted. The findings were shown in Table 4.5.

**Table 4.5. Initial Gain in Receptive Word Knowledge at Lower-Intermediate Level**

|                  | Groups | N  | $\bar{x}$ | $X^2$  | p        |
|------------------|--------|----|-----------|--------|----------|
| <b>Immediate</b> | SR     | 21 | 7         | 58.164 | < 0.001* |
| <b>Receptive</b> | SW     | 18 | 7         |        |          |
| <b>Post-test</b> | FR     | 20 | 6.25      |        |          |
|                  | FW     | 20 | 3.75      |        |          |
|                  | RO     | 21 | 2         |        |          |
|                  | WO     | 23 | 4         |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

With respect to receptive word knowledge, it can be seen that all the groups, regardless of doing variously loaded tasks, gained at least a few of the eight target words. When all the participants' immediate receptive vocabulary post-test scores were compared, it was found

that the learners significantly differed in their initial gain of the correct meanings ( $p < 0.001$ ). When the level of task-induced involvement load was taken into account, it was noticed that both sentence-writing groups (index = 3) significantly outsourced the FW (index = 2) and control groups (index = 0) and did moderately better than the FR group (index = 2) on the immediate receptive post-test. As anticipated, these learners who performed the SR and SW tasks with the highest degree of involvement (index = 3) acquired the greatest vocabulary gain. The FR group also significantly outperformed the control groups. Therefore, it may be proposed that tasks that induce a greater degree of involvement can result in more vocabulary gain than those that have a lower level of task involvement. The only contradictory finding to this assumption was the average performance of the WO group. Interestingly, this group also performed considerably well on the immediate receptive post-test despite not tackling with the target words during their task.

#### 4.2.1.5. Retention in Receptive Word Knowledge

The results of a Kruskal Wallis test indicated the extent to which the lower-intermediate learners were able to recall the eight target words two weeks after their task implementation. Table 4.6 shows the learners' receptive vocabulary retention on the second measurement.

**Table 4.6. Retention in Receptive Word Knowledge at Lower-Intermediate Level**

|                  | Groups | <i>N</i> | $\bar{x}$ | $X^2$  | <i>p</i> |
|------------------|--------|----------|-----------|--------|----------|
| <b>Delayed</b>   | SR     | 21       | 7         | 57.753 | < 0.001* |
| <b>Receptive</b> | SW     | 18       | 7.5       |        |          |
| <b>Post-test</b> | FR     | 20       | 6.25      |        |          |
|                  | FW     | 20       | 2         |        |          |
|                  | RO     | 21       | 2         |        |          |
|                  | WO     | 23       | 3         |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

As the results indicated, there was a statistically significant difference among the groups ( $p < 0.001$ ). The SW, SR and FR groups retrieved more target words by significantly outperforming the remaining groups. In other words, the FW, RO and WO groups scored

significantly lower on the delayed receptive post-test. These findings showed that the learners who completed the tasks with a higher involvement load attained greater retention in meaning over time than those who were assigned to the tasks with a lower degree of involvement load. This finding was supported by the outcome of both sentence writing tasks (index = 3). Likewise, the FR task (index = 2) also led to more retention in meaning than the control group tasks (index = 0) because it had a higher index of involvement load. However, the FW task (index = 2) did not exhibit the same pattern. The learners who performed this task did not recall more target words than the learners in the control groups.

#### **4.2.1.6. Changes in Receptive Word Knowledge within Groups**

To find out whether the lower-intermediate learners' receptive word knowledge significantly increased or dropped over time, a Wilcoxon signed-ranked test was conducted (for the results of the test, see Appendix L). The comparison of the immediate and delayed receptive vocabulary post-test scores indicated a significant change within one group only. The FW group significantly lost their knowledge of the target words within two weeks' time ( $Z = -2.437$ ,  $p = 0.015$ ). Surprisingly, the remaining groups recalled the target words with less difficulty.

#### **4.2.2. Vocabulary Gain from Tasks with Written vs. Audiovisual Input Mode**

The second research question investigated whether vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at lower-intermediate proficiency level in terms of productive and receptive word knowledge. In order to answer this question, Kruskal Wallis tests were conducted by comparing the immediate and delayed productive and receptive vocabulary post-test scores of the lower-intermediate learners who performed equally loaded tasks through different kinds of input. The results were presented in the following sections.

##### **4.2.2.1. Gain and Retention in Productive Word Knowledge**

The immediate and delayed productive vocabulary post-test scores of the lower-intermediate participants who were assigned to the tasks with an identical involvement index but different types of input mode were shown in Table 4.7.

**Table 4.7. Productive Word Gain and Retention at Lower-Intermediate Level (Input Based)**

| Tasks             |           | Control<br>(index = 0) |                         | Fill-in<br>(index = 2) |                         | Sentence Writing<br>(index = 3) |                         |
|-------------------|-----------|------------------------|-------------------------|------------------------|-------------------------|---------------------------------|-------------------------|
| Input Mode        |           | Written<br>(N = 21)    | Audiovisual<br>(N = 23) | Written<br>(N = 20)    | Audiovisual<br>(N = 20) | Written<br>(N = 21)             | Audiovisual<br>(N = 18) |
| <b>Immediate</b>  | $\bar{x}$ | 0.75                   | 0                       | 5.38                   | 2                       | 5.5                             | 4                       |
| <b>Productive</b> | Min.      | 0                      | 0                       | 1                      | 0                       | 1.25                            | 1                       |
| <b>Post-test</b>  | Max.      | 1                      | 3.75                    | 8                      | 7.75                    | 7.5                             | 7.75                    |
| <b>Delayed</b>    | $\bar{x}$ | 0.25                   | 1                       | 3                      | 0                       | 2                               | 3.38                    |
| <b>Productive</b> | Min.      | 0                      | 0                       | 1                      | 0                       | 0                               | 0                       |
| <b>Post-test</b>  | Max.      | 3                      | 3                       | 7.75                   | 4.5                     | 7.5                             | 8                       |

**Note.** Written: Reading a passage as the medium of teaching, Audiovisual: Watching a video podcast as the medium of teaching. Scores were given out of 8 points.

Table 4.7 reveals that when the immediate productive vocabulary post-test scores of the groups were compared based on the kind of input they were subjected to, it can be seen that the learners who performed their tasks through the medium of reading (SR, FR and RO) attained greater initial productive word knowledge than their counterparts who completed the same tasks through the medium of watching (SW, FW and WO). In other words, the learners who were assigned to the written input performed better than their peers who were exposed to the audiovisual input through the same task, but no significant differences were found within each task type. Therefore, it can be concluded that tasks implemented through the written input lead to more knowledge of the target forms in the short run than those implemented through the audiovisual input.

As shown in Table 4.7, the results of the delayed productive vocabulary post-test indicated a change in the efficacy of the written input over time. Only the FR group significantly recalled more target forms than their correspondents in the FW group, but the audiovisual input led to slightly more productive word retention among the sentence writing and control groups. This shows that the audiovisual input can yield more permanent vocabulary gain than the written input in the long run.

When the input type, the task index and the time effect were taken into account together, it was found for the lower-intermediate learners that the written input ultimately leads to more productive word knowledge when it is presented through a gap filling task (index = 2), while the audiovisual input leads to more productive word knowledge over time when it

is presented through a sentence writing task (index = 3). In relation to the aforementioned results, it can be concluded that sentence writing tasks seem to help lower-proficiency learners acquire the correct forms when they are performed by means of an audiovisual resource, whereas gap filling tasks enhance more productive word knowledge when they are completed along with a written text. Finally, as it can be anticipated, the control group tasks (index = 0) produced the smallest immediate vocabulary gain of the correct forms regardless of the input mode. Following the FW group, they still retrieved the fewest target forms on the delayed post-test.

#### 4.2.2.2. Gain and Retention in Receptive Word Knowledge

The immediate and delayed receptive vocabulary post-test scores of the lower-intermediate participants who were instructed to do the tasks with the same involvement index but two kinds of input mode were presented in Table 4.8.

**Table 4.8. Receptive Word Gain and Retention at Lower-Intermediate Level (Input Based)**

| Tasks            |           | Control<br>(index = 0) |                         | Fill-in<br>(index = 2) |                         | Sentence Writing<br>(index = 3) |                         |
|------------------|-----------|------------------------|-------------------------|------------------------|-------------------------|---------------------------------|-------------------------|
| Input Mode       |           | Written<br>(N = 21)    | Audiovisual<br>(N = 23) | Written<br>(N = 20)    | Audiovisual<br>(N = 20) | Written<br>(N = 21)             | Audiovisual<br>(N = 18) |
| <b>Immediate</b> | $\bar{x}$ | 2                      | 4                       | 6.25                   | 3.75                    | 7                               | 7                       |
| <b>Receptive</b> | Min.      | 0                      | 0                       | 2                      | 0                       | 3                               | 3                       |
| <b>Post-test</b> | Max.      | 6                      | 6                       | 8                      | 8                       | 8                               | 8                       |
| <b>Delayed</b>   | $\bar{x}$ | 2                      | 3                       | 6.25                   | 2                       | 7                               | 7.5                     |
| <b>Receptive</b> | Min.      | 0                      | 1                       | 0                      | 0                       | 4                               | 2                       |
| <b>Post-test</b> | Max.      | 6                      | 6                       | 8                      | 8                       | 8                               | 8                       |

**Note.** Written: Reading a passage as the medium of teaching, Audiovisual: Watching a video podcast as the medium of teaching. Scores were given out of 8 points.

Table 4.8 reveals that when the immediate receptive vocabulary post-test scores of the lower-intermediate learners were compared, it can be seen that the sentence writing groups exhibited the same performance regardless of the input type. However, while the written input seemed to be more effective in enhancing more initial knowledge of meaning within the fill-in tasks, the audiovisual input yielded more receptive word knowledge in the short term among the control groups. The results of the Kruskal Wallis test comparing the delayed receptive vocabulary post-test scores reached at the same findings, but this time the difference between the FR ( $Mdn = 6.25$ ) and FW ( $Mdn = 2$ ) groups turned out to be significant ( $H(5) = 57.753, p < 0.001$ ).

Regarding the link between input modality and task type, the aforementioned results revealed that the audiovisual input enhanced slightly more retention of the target words among the sentence writing and control groups, while the written input made a bigger difference within the gap filling tasks. With respect to the effects of input mode on task-induced involvement load, it can be concluded that the lower-intermediate learners recall a greater number of target words when they are exposed to a written text through a gap filling task (index = 2), whereas with a sentence writing task (index = 3) input type seems to play a relatively small role in vocabulary acquisition.

On the other hand, both receptive vocabulary post-tests showed that sentence writing tasks (index = 3) foster more vocabulary gain and retention among lower-proficiency learners than fill-in tasks (index = 2) regardless of input type. It was seen that the learners who formed sentences by incorporating the target items were able to acquire and recall these items better than those who performed a gap filling task. However, the same results revealed that the FW task (index = 2) led to a similar amount of vocabulary gain and retention to those of the control group tasks (index = 0) although the former task had a higher level of involvement load.

### **4.3. Results for Upper-Intermediate Proficiency Level**

The results associated with the upper-intermediate participants were presented in the subsequent parts of this chapter by addressing the related research questions.

#### **4.3.1. Vocabulary Gain from Tasks with Low vs. High Involvement Load**

The third research question aimed to explore whether vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at upper-intermediate proficiency level in terms of productive and receptive word knowledge. The upper-intermediate learners' median scores on the immediate and delayed productive and receptive vocabulary post-tests were compared using Kruskal Wallis tests to detect the significant differences among the learners who performed variously loaded tasks through two input modes. The results of these tests were shown in the following sections.

#### 4.3.1.1. Gain in Productive Word Knowledge

The median scores of the upper-intermediate learners on the immediate and delayed productive vocabulary post-tests were provided in Table 4.9.

**Table 4.9. Productive Word Gain and Retention at Upper-Intermediate Level (Index Based)**

| Groups | Index | N  | Immediate   |             | Delayed     |             |
|--------|-------|----|-------------|-------------|-------------|-------------|
|        |       |    | $\tilde{x}$ | Min. – Max. | $\tilde{X}$ | Min. – Max. |
| SR     | 3     | 19 | 3.75        | 0 – 8       | 2           | 0 – 4.5     |
| SW     | 3     | 17 | 4.75        | 0.25 – 8    | 2.75        | 0 – 6       |
| FR     | 2     | 17 | 3           | 0 – 7.5     | 1.75        | 0 – 3       |
| FW     | 2     | 21 | 4.75        | 0 – 8       | 1.5         | 0 – 7       |
| RO     | 0     | 19 | 0           | 0 – 3       | 1           | 0 – 3       |
| WO     | 0     | 20 | 0           | 0 – 2       | 1           | 0 – 4       |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only. Scores were given out of 8 points.

In order to determine whether the learners who were subjected to the eight target words through six variously loaded tasks differed in terms of their immediate productive word gain, a Kruskal Wallis test was used to compare their median scores on the immediate productive vocabulary post-test. Table 4.10 shows this comparison.

**Table 4.10. Initial Gain in Productive Word Knowledge at Upper-Intermediate Level**

|                   | Groups | N  | $\tilde{x}$ | $X^2$  | p        |
|-------------------|--------|----|-------------|--------|----------|
| <b>Immediate</b>  | SR     | 19 | 3.75        | 66.841 | < 0.001* |
| <b>Productive</b> | SW     | 17 | 4.75        |        |          |
| <b>Post-test</b>  | FR     | 17 | 3           |        |          |
|                   | FW     | 21 | 4.75        |        |          |
|                   | RO     | 19 | 0           |        |          |
|                   | WO     | 20 | 0           |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

Table 4.10 demonstrates that there was a statistically significant difference among the groups in their immediate productive vocabulary post-test scores ( $p < 0.001$ ). It was seen that the control groups significantly obtained lower scores than the remaining groups. However, the experimental groups did not significantly differ in their immediate productive

word gain. With respect to the order of the medians, the SW and FW groups scored the highest ( $Mdn = 4.75$ ) followed by the SR ( $Mdn = 3.75$ ) and FR ( $Mdn = 3$ ) groups.

Based on the results presented above, the groups were compared by looking at the involvement load degrees of their tasks to determine which tasks lead to more productive word gain. It was found that while the SW and FW groups scored the same, the SR group slightly outperformed the FR group. This finding pointed out the superiority of sentence writing tasks (index = 3) over gap filling tasks (index = 2) to some extent. It was further supported by the zero knowledge of the target forms that the control groups acquired since their tasks had no involvement load.

#### 4.3.1.2. Retention in Productive Word Knowledge

Two weeks after the task implementation, the upper-intermediate learners were administered with the delayed productive vocabulary post-test to find out whether they were still able to recall the eight target forms. The participants' scores on the delayed productive vocabulary post-test were analyzed using a Kruskal Wallis test. Table 4.11 presents the results of each task group.

**Table 4.11. Retention in Productive Word Knowledge at Upper-Intermediate Level**

|                   | Groups | N  | $\bar{x}$ | $X^2$  | p        |
|-------------------|--------|----|-----------|--------|----------|
| <b>Delayed</b>    | SR     | 19 | 2         | 23.765 | < 0.001* |
| <b>Productive</b> | SW     | 17 | 2.75      |        |          |
| <b>Post-test</b>  | FR     | 17 | 1.75      |        |          |
|                   | FW     | 21 | 1.5       |        |          |
|                   | RO     | 19 | 1         |        |          |
|                   | WO     | 20 | 1         |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

The results of the Kruskal Wallis test revealed a statistically significant difference among the groups ( $p < 0.001$ ). When productive word knowledge was measured again, it was found that the SW group ( $Mdn = 2.75$ ) slightly outsourced all the other groups, but their performance turned out to be significant only compared to the scores of the control groups.

In other words, the only statistically significant difference was observed between the SW and control groups.

Considering that the highest scores on the delayed productive vocabulary post-test were obtained through the sentence writing tasks (index = 3) and the learners who performed these tasks recalled the correct forms better than those assigned to the gap filling tasks (index = 2), it can be suggested that tasks with a higher degree of involvement load lead to more retention of form. This finding seemed to apply to the gap filling groups too because their tasks also enhanced more productive word retention than the RO and WO tasks (index = 0).

#### **4.3.1.3. Changes in Productive Word Knowledge within Groups**

In order to compare the immediate and delayed productive vocabulary post-test scores of the upper-intermediate learners to find out whether they preserved or lost their knowledge of the target forms over time, a Wilcoxon signed-ranked test was conducted (for the results of the test, see Appendix M). The results demonstrated that all the groups experienced a significant change within two weeks following the task implementation. The learners in the SW group who achieved the highest score on the immediate post-test significantly lost their knowledge of the correct forms over time ( $Z = -3.623, p < 0.001$ ). Similarly, the scores of the FW group who also obtained the highest score on the immediate post-test significantly decreased on the delayed post-test ( $Z = -2.614, p = 0.009$ ). The SR ( $Z = -2.702, p = 0.007$ ) and FR ( $Z = -3.424, p = 0.001$ ) groups also had a significant loss in their productive word knowledge. Surprisingly, the WO ( $Z = -2.812, p = 0.005$ ) and RO ( $Z = -2.470, p = 0.014$ ) groups gained significantly more productive word knowledge over time.

#### **4.3.1.4. Gain in Receptive Word Knowledge**

The median scores of the upper-intermediate learners on the immediate and delayed receptive vocabulary post-tests were given in Table 4.12.

**Table 4.12. Receptive Word Gain and Retention at Upper-Intermediate Level (Index Based)**

| Groups | Index | N  | Immediate |             | Delayed   |             |
|--------|-------|----|-----------|-------------|-----------|-------------|
|        |       |    | $\bar{x}$ | Min. – Max. | $\bar{x}$ | Min. – Max. |
| SR     | 3     | 19 | 6         | 1 – 8       | 5         | 1.5 – 8     |
| SW     | 3     | 17 | 7         | 3 – 8       | 5         | 1 – 7       |
| FR     | 2     | 17 | 4         | 1 – 7       | 3         | 0 – 6       |
| FW     | 2     | 21 | 6         | 1 – 8       | 3.5       | 0 – 6.5     |
| RO     | 0     | 19 | 1         | 0 – 6       | 2         | 0 – 5       |
| WO     | 0     | 20 | 2         | 0 – 5       | 2.5       | 0 – 7       |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only. Scores were given out of 8 points.

A Kruskal Wallis test was conducted to see whether there were any significant differences among the upper-intermediate learners in their immediate receptive vocabulary post-test scores. The findings were provided in Table 4.13.

**Table 4.13. Initial Gain in Receptive Word Knowledge at Upper-Intermediate Level**

| Groups                               |    | N  | $\bar{x}$ | X <sup>2</sup> | p        |
|--------------------------------------|----|----|-----------|----------------|----------|
| <b>Immediate Receptive Post-test</b> | SR | 19 | 6         | 66.855         | < 0.001* |
|                                      | SW | 17 | 7         |                |          |
|                                      | FR | 17 | 4         |                |          |
|                                      | FW | 21 | 6         |                |          |
|                                      | RO | 19 | 1         |                |          |
|                                      | WO | 20 | 2         |                |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

It can be deduced from Table 4.13 that all the groups, regardless of performing different tasks, acquired at least a few of the eight target words. When the participants' immediate receptive vocabulary post-test scores were analyzed, it was found that the groups significantly differed in their initial gain of the correct meanings ( $p < 0.001$ ). The highest median score was achieved by the SW group ( $Mdn = 7$ ) who significantly outperformed the learners in the control groups. Following these learners, the SR and FW groups shared the same score ( $Mdn = 6$ ) and also significantly did better than the control groups. The median scores of the aforementioned groups were really approximate to the total number of the

target words. This implies that the learners in these groups acquired most of the target words. The last significant difference was detected between the FR ( $Mdn = 4$ ) and RO ( $Mdn = 1$ ) groups. Finally, it was unexpected, but the control groups also gained some receptive word knowledge as a result of performing their tasks.

When the degree of task-induced involvement load was considered, it was seen that both sentence writing groups (index = 3) outscored their corresponding gap filling groups (index = 2). In other words, the SW group did better on the immediate productive vocabulary post-test than the FW group. Similarly, the SR group outperformed the FR group on the same test. The gap filling groups also obtained higher scores than the control groups (index = 0). As a result, it can be assumed that tasks that induce a higher level of involvement can initially enhance more receptive word knowledge than those that have a lower level of task involvement.

#### 4.3.1.5. Retention in Receptive Word Knowledge

The results of a Kruskal Wallis test showed the extent to which the upper-intermediate learners retrieved the eight target words within two weeks' time. Table 4.14 presents the learners' receptive vocabulary retention on the second testing procedure.

**Table 4.14. Retention in Receptive Word Knowledge at Upper-Intermediate Level**

|                  | Groups | <i>N</i> | $\tilde{x}$ | $X^2$  | <i>p</i> |
|------------------|--------|----------|-------------|--------|----------|
| <b>Delayed</b>   | SR     | 19       | 5           | 31.863 | < 0.001* |
| <b>Receptive</b> | SW     | 17       | 5           |        |          |
| <b>Post-test</b> | FR     | 17       | 3           |        |          |
|                  | FW     | 21       | 3.5         |        |          |
|                  | RO     | 19       | 2           |        |          |
|                  | WO     | 20       | 2.5         |        |          |

**Note.** SR: Sentence-writing-by-reading (index = 3), SW: Sentence-writing-by-watching (index = 3), FR: Fill-in-by-reading (index = 2), FW: Fill-in-by-watching (index = 2), RO: Reading-only (index = 0), WO: Watching-only (index = 0), \* $p < .05$ . Scores were given out of 8 points.

The results indicated that there was a statistically significant difference among the groups ( $p < 0.001$ ). Regarding the order of the median scores, it can be seen that the SW and SR groups both obtained the highest score and significantly achieved better than the control

groups. The FW and FR groups also outscored the control groups, but the difference was not significant. The only significant difference among the experimental groups was between the SR and FR groups.

The results of the delayed receptive vocabulary post-test pointed out that the greatest word gain was obtained through the sentence writing tasks (index = 3). The upper-intermediate learners who did the tasks with a higher degree of involvement load recalled more target words than those who were assigned to the tasks with a lower degree of involvement load. The significant superiority of the SR group over the FR group highlighted this finding. Likewise, the gap filling tasks (index = 2) also led to more retention in meaning than the control group tasks (index = 0).

#### **4.3.1.6. Changes in Receptive Word Knowledge within Groups**

To determine the changes in the upper-intermediate learners' receptive word knowledge over a period of two weeks, a Wilcoxon signed-ranked test was conducted. The results revealed a significant change in receptive word knowledge within three groups only. The SW group significantly lost their knowledge of the target items over time ( $Z = -3.319$ ,  $p = 0.001$ ). Similarly, the immediate receptive vocabulary post-test scores of the FW ( $Z = -2.473$ ,  $p = 0.013$ ) and FR ( $Z = -2.466$ ,  $p = 0.014$ ) groups also dropped on the delayed receptive vocabulary post-test.

#### **4.3.2. Vocabulary Gain from Tasks with Written vs. Audiovisual Input Mode**

The last research question investigated whether vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at upper-intermediate proficiency level in terms of productive and receptive word knowledge. In order to explore this question, Kruskal Wallis tests were carried out by comparing the immediate and delayed productive and receptive vocabulary post-test scores of the upper-intermediate learners who completed identically loaded tasks through two types of input. The findings were provided in the following sections.

#### 4.3.2.1. Gain and Retention in Productive Word Knowledge

The immediate and delayed productive vocabulary post-test scores of the upper-intermediate participants who were instructed to do the tasks with the same involvement index but different kinds of input mode were demonstrated in Table 4.15.

**Table 4.15. Productive Word Gain and Retention at Upper-Intermediate Level (Input Based)**

| Tasks             |           | Control<br>(index = 0) |                         | Fill-in<br>(index = 2) |                         | Sentence Writing<br>(index = 3) |                         |
|-------------------|-----------|------------------------|-------------------------|------------------------|-------------------------|---------------------------------|-------------------------|
| Input Mode        |           | Written<br>(N = 19)    | Audiovisual<br>(N = 17) | Written<br>(N = 17)    | Audiovisual<br>(N = 21) | Written<br>(N = 19)             | Audiovisual<br>(N = 20) |
| <b>Immediate</b>  | $\bar{x}$ | 0                      | 0                       | 3                      | 4.75                    | 3.75                            | 4.75                    |
| <b>Productive</b> | Min.      | 0                      | 0                       | 0                      | 0                       | 0                               | 0.25                    |
| <b>Post-test</b>  | Max.      | 3                      | 2                       | 7.5                    | 8                       | 8                               | 8                       |
| <b>Delayed</b>    | $\bar{x}$ | 1                      | 1                       | 1.75                   | 1.5                     | 2                               | 2.75                    |
| <b>Productive</b> | Min.      | 0                      | 0                       | 0                      | 0                       | 0                               | 0                       |
| <b>Post-test</b>  | Max.      | 3                      | 4                       | 3                      | 7                       | 4.5                             | 6                       |

**Note.** Written: Reading a passage as the medium of teaching, Audiovisual: Watching a video podcast as the medium of teaching. Scores were given out of 8 points.

As shown in Table 4.15, when the experimental tasks were compared in terms of the input they presented, the results revealed that the learners who completed their tasks through the medium of watching (SW and FW) attained greater initial productive word knowledge than those who performed the same tasks through the medium of reading (SR and FR). To put it simply, both types of tasks were equally most successful when they were completed through the audiovisual input. Therefore, it can be concluded that the use of audiovisual materials is more likely to result in the acquisition of the target forms than the use of written texts in the short term. As for the control groups, no difference was identified regarding the effects of the two input modalities.

The results of the delayed productive vocabulary post-test indicated similar findings for the upper-intermediate learners. When the groups were compared based on the input they received, it was found that the audiovisual input led to slightly more retention of form within the sentence writing groups, while being exposed to the written input during the gap filling tasks enhanced more productive word knowledge in the long run. Finally, the control groups indicated no difference concerning the influence of input modality on productive word retention.

#### 4.3.2.2. Gain and Retention in Receptive Word Knowledge

The immediate and delayed receptive vocabulary post-test scores of the upper-intermediate participants who were instructed to do the tasks with the same involvement index but two kinds of input mode were presented in Table 4.16.

**Table 4.16. Receptive Word Gain and Retention at Upper-Intermediate Level (Input Based)**

| Tasks      |           | Control<br>(index = 0) |                         | Fill-in<br>(index = 2) |                         | Sentence Writing<br>(index = 3) |                         |
|------------|-----------|------------------------|-------------------------|------------------------|-------------------------|---------------------------------|-------------------------|
| Input Mode |           | Written<br>(N = 19)    | Audiovisual<br>(N = 17) | Written<br>(N = 17)    | Audiovisual<br>(N = 21) | Written<br>(N = 19)             | Audiovisual<br>(N = 20) |
| Immediate  | $\bar{x}$ | 1                      | 2                       | 4                      | 6                       | 6                               | 7                       |
| Receptive  | Min.      | 0                      | 0                       | 1                      | 1                       | 1                               | 3                       |
| Post-test  | Max.      | 6                      | 5                       | 7                      | 8                       | 8                               | 8                       |
| Delayed    | $\bar{x}$ | 2                      | 2.5                     | 3                      | 3.5                     | 5                               | 5                       |
| Receptive  | Min.      | 0                      | 0                       | 0                      | 0                       | 1.5                             | 1                       |
| Post-test  | Max.      | 5                      | 7                       | 6                      | 6.5                     | 8                               | 7                       |

**Note.** Written: Reading a passage as the medium of teaching, Audiovisual: Watching a video podcast as the medium of teaching. Scores were given out of 8 points.

The results of the immediate receptive vocabulary post-test indicated that although no significant differences were noticed between the learners who completed the same tasks involving different types of input, the SW, FW and WO groups provided more accurate definitions for the target items than the SR, FR and RO groups respectively. In other words, all the task groups initially acquired more target words through the audiovisual input than the written input. Therefore, it can be concluded that tasks implemented through the audiovisual input lead to more vocabulary gain in the short term than tasks implemented through the written input.

When input type was taken into account again, the results of the delayed receptive vocabulary post-test showed that if presented through gap filling tasks, the audiovisual input fosters more receptive word retention among higher-proficiency learners. Therefore, it can be concluded that gap filling tasks lead to more retention of the target words when performed through the audiovisual input rather than the written input. However, this finding should be interpreted cautiously since the difference between the FW and FR groups seemed to be rather small (i.e., 0.5 word) to make a valid generalization. On the other hand, the audiovisual input did not create any difference within sentence writing tasks. The upper-intermediate learners retrieved an equal number of target words when they

were shown an audiovisual resource or provided with a written text before doing a sentence writing task.

#### 4.4. Summary of the Results

The results of the vocabulary post-tests showed that the findings regarding variously loaded tasks were largely consistent in terms of the effect of task-induced involvement load on incidental vocabulary learning. However, the findings concerning identically loaded tasks were somewhat contradictory when input modality interacted with L2 proficiency. The summary of these findings were provided below.

**Table 4.17. Summary of the Findings**

| <b>PRODUCTIVE WORD KNOWLEDGE</b>                                                                                                                                                                                                                                                             |                                                                                                                                                                                                                                                                                                                                            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>LOWER-INTERMEDIATE</b>                                                                                                                                                                                                                                                                    | <b>UPPER-INTERMEDIATE</b>                                                                                                                                                                                                                                                                                                                  |
| <b>SHORT-TERM</b>                                                                                                                                                                                                                                                                            | <b>SHORT-TERM</b>                                                                                                                                                                                                                                                                                                                          |
| <ul style="list-style-type: none"> <li>• Sentence writing is more effective than gap filling only with the same input.</li> <li>• The written input is more effective than the audiovisual input regardless of task type.</li> </ul>                                                         | <ul style="list-style-type: none"> <li>• Gap filling and sentence writing are equally effective only with the audiovisual input.</li> <li>• Sentence writing is more effective than gap filling only with the written input.</li> <li>• The audiovisual input is more effective than the written input regardless of task type.</li> </ul> |
| <b>LONG-TERM</b>                                                                                                                                                                                                                                                                             | <b>LONG-TERM</b>                                                                                                                                                                                                                                                                                                                           |
| <ul style="list-style-type: none"> <li>• Gap filling is more effective than sentence writing only with the written input.</li> <li>• Sentence writing is more effective than gap filling only with the audiovisual input.</li> </ul>                                                         | <ul style="list-style-type: none"> <li>• Sentence writing is more effective than gap filling regardless of input mode.</li> </ul>                                                                                                                                                                                                          |
| <b>BOTH LEVELS</b>                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                            |
| <ul style="list-style-type: none"> <li>• The paring of gap filling with the written input and sentence writing with the audiovisual input are both effective in the long term.</li> </ul>                                                                                                    |                                                                                                                                                                                                                                                                                                                                            |
| <b>RECEPTIVE WORD KNOWLEDGE</b>                                                                                                                                                                                                                                                              |                                                                                                                                                                                                                                                                                                                                            |
| <b>LOWER-INTERMEDIATE</b>                                                                                                                                                                                                                                                                    | <b>UPPER-INTERMEDIATE</b>                                                                                                                                                                                                                                                                                                                  |
| <b>SHORT-TERM</b>                                                                                                                                                                                                                                                                            | <b>SHORT-TERM</b>                                                                                                                                                                                                                                                                                                                          |
| <ul style="list-style-type: none"> <li>• Sentence writing is more effective than gap filling regardless of input mode.</li> <li>• The written input is more effective than the audiovisual input with gap filling.</li> <li>• There is no input difference with sentence writing.</li> </ul> | <ul style="list-style-type: none"> <li>• Sentence writing is more effective than gap filling only with the same input.</li> <li>• The audiovisual input is more effective than the written input regardless of task type.</li> </ul>                                                                                                       |
| <b>LONG-TERM</b>                                                                                                                                                                                                                                                                             | <b>LONG-TERM</b>                                                                                                                                                                                                                                                                                                                           |
| <ul style="list-style-type: none"> <li>• The written input is more effective than the audiovisual input with gap filling.</li> <li>• The audiovisual input is more effective than the written input with sentence writing.</li> </ul>                                                        | <ul style="list-style-type: none"> <li>• The audiovisual input is more effective than the written input with gap filling.</li> <li>• There is no input difference with sentence writing.</li> </ul>                                                                                                                                        |
| <b>BOTH LEVELS</b>                                                                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                            |
| <ul style="list-style-type: none"> <li>• Sentence writing is more effective than gap filling in the long term regardless of input mode.</li> </ul>                                                                                                                                           |                                                                                                                                                                                                                                                                                                                                            |

## **CHAPTER V: DISCUSSION AND CONCLUSION**

In this chapter, the findings of the present study are discussed in relation to the research questions addressed at the beginning of the study. The findings are also reviewed by referring to the assumptions of the Involvement Load Hypothesis and results of the relevant literature. The pedagogical implications and limitations of the study and suggestions for future research are also provided in this section.

### **5.1. Discussion of the Findings**

The current study attempted to investigate the effects of the construct of task-induced involvement load, input type and L2 proficiency on incidental vocabulary learning. Six different tasks (primarily including reading/watching-only, gap filling and sentence writing) with varying indexes of involvement load (0, 2 or 3) and two kinds of input mode (written or audiovisual) were created to be implemented in the study. Data were gathered through productive and receptive vocabulary post-tests administered immediately after the task implementation and two weeks later in 2016-2017 academic year. The results concerning each research question of the study are discussed in the following parts of the chapter.

#### **5.1.1. Vocabulary Gain from Various Loaded Tasks at Lower-Intermediate Level**

The first research question of the current study aimed to explore whether vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at lower-intermediate proficiency level in terms of productive and receptive word knowledge. To answer this research question, three kinds of vocabulary tasks were formed for the lower-proficiency group. These tasks mainly involved a reading/watching-only task, a gap filling task and a sentence writing task. The tasks consisted of an index of 0, 2 and 3 respectively. The gap filling and sentence writing tasks were designed to share the same degree of the search component but a varying level of evaluation, while the control group tasks had neither search nor evaluation. This task design would help test the effects of various loaded tasks on incidental vocabulary acquisition,

whereas the presence of the control groups would help better understand the importance of the task-induced involvement load in L2 vocabulary learning.

For productive word knowledge, the results of the lower-intermediate learners indicated that sentence writing tasks (index = 3) yield more initial knowledge of the target forms than fill-in tasks (index = 2) if these tasks are performed through the same kind of input (i.e., written or audiovisual). Regardless of input type, both types of tasks enhance more productive word knowledge than reading/watching-only tasks with no involvement load (index = 0). For the retention of the correct forms in the long term, the superiority of sentence writing tasks is still maintained only through the audiovisual input, while gap filling tasks seem to be almost equally effective when they are completed through the written input. This means that productive word knowledge is fostered more when L2 learners have to write original sentences with the target items. In other words, the attempt to write semantically appropriate sentences with the target words might encourage the learners to acquire them. This can also be explained by the strong degree of the evaluation component induced by sentence writing tasks. As Kim (2011) claimed, the three components of the Involvement Load Hypothesis may not be equal in value. Strong evaluation can contribute to vocabulary enhancement more because it generates more cognitive effort and thus facilitates retention.

In addition to sentence writing tasks, filling the blanks left out in a reading passage with the target words can help lower-proficiency learners attain a good amount of both productive and receptive word knowledge as revealed in this study with the efficacy of the FR task. To examine this issue, several justifications can be made. First, the learners in this study were accustomed to doing fill-in-the-blank exercises due to the instructional approach of their course book and layout of the vocabulary questions on their examinations. As a result, they might be inclined to commit the target forms into memory while supplying them in the gapped reading. Moreover, the focus on writing the target forms correctly during the task might have prompted these learners to acquire the form as well as the meaning.

On the other hand, it should be noted here that despite having a high degree of involvement load, the fill-in-by-watching task (index = 2) did not create the same effect as the FR task (index = 2) did. Contrary to the predictions of the Involvement Load Hypothesis, the FW group scored lower than almost all the other groups on both productive and receptive post-tests. This is a striking result because these learners were supposed to outperform the control groups at least owing to the higher involvement index of their task. This unexpected finding can be attributed to the poor listening skills possessed by the lower-intermediate group. That is why the learners in the FW group struggled while keeping track of the flow of speech in the video podcast, so they were unable to identify the target words most of the time. They lacked the necessary bottom-up skills to recognize and decode individual words. As a result, they could not supply most of the correct forms and meanings on both productive and receptive post-tests.

With respect to receptive vocabulary gain and retention of the lower-intermediate participants, it was found that both sentence writing tasks (index = 3) yielded a greater amount of meaning recall than gap filling tasks (index = 2). This finding is consistent with those of the earlier studies conducted by Hulstijn and Laufer (2001), Kim (2011), Mármol and Sánchez-Lafuente (2013) and Tu (2004). They similarly discovered that the more involving task (sentence writing) leads to more passive recall of meaning than the less involving task (gap filling). Therefore, as Kim (2011) suggested, although involving tasks may be cognitively challenging for less proficient L2 learners due to their limited language abilities, they can still be favoured by these learners as well. Similarly, Swain's (1985) Output Hypothesis stated that the act of producing language makes L2 learners become aware of the gap in their linguistic knowledge and try to compensate for it to express their intended message clearly. Thus, it is obvious why the sentence writing groups reflected on the definitions of the target words more. They wanted to convey an accurate and meaningful message with their sentences. This helped them recall the target words with less difficulty and score higher than the gap filling groups on both receptive post-tests.

When reading-only and watching-only tasks are considered, it can be inferred that such tasks play a relatively insignificant role in facilitating the learning and retrieval of the

correct forms and meanings. Fully verifying the Involvement Load Hypothesis, these tasks resulted in a rather low amount of productive word knowledge and a slightly moderate amount of receptive word knowledge. This was because the lower-intermediate learners assigned to these tasks were not directly exposed to the target words by means of any enhancing effects such as highlighting or glossing, so they did not notice these items. These learners only performed a reading/video comprehension task. That is why they did not acquire most of the new words. As Schmidt (1990) suggested, L2 learners need to notice the new information to be able to learn it. Likewise, Laufer (2003) stated that explicit focus through vocabulary tasks is indispensable for the deep processing and learning of form-meaning associations. This explains why the learners who dealt with the target words through the sentence writing or gap filling tasks were able to outperform the others who only read a passage or watched a video podcast with the target items not highlighted or glossed. This conclusion emphasized that vocabulary tasks with higher degrees of involvement load appear to be more productive in lexical acquisition than mere reading (or video comprehension as shown in this study).

Partially confirming the assumptions of the Involvement Load Hypothesis, the findings of the lower-intermediate participants were mainly in harmony with those of the previous studies carried out by Hulstijn and Laufer (2001), Laufer (2003), Llach (2009), Lu (2013), Keating (2008), Kim (2011) and Tu (2004) in that the most involving task (sentence writing) enhanced more vocabulary gain and retention than the more involving task (gap filling) which in turn led to more vocabulary gain and retention than the least involving task (reading or video comprehension).

### **5.1.2. Vocabulary Gain from Different Modes of Input at Lower-Intermediate Level**

The second research question of the present study investigated whether vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at lower-intermediate proficiency level in terms of productive and receptive word knowledge. To answer this research question, three types of vocabulary tasks (sentence writing, gap filling and control group tasks) were each combined with the two varieties of input mode (written and audiovisual). In this way, each task was divided into two different

versions (one with the written input and the other with the audiovisual input) that shared an identical index of involvement load. This specific task design would help test three assumptions relating to the Involvement Load Hypothesis. First, it would be easier to distinguish between the effectiveness of the audiovisual input and the written input in vocabulary acquisition in general. Second, this task layout would empirically test the prediction of the hypothesis that tasks with an equal index of involvement load lead to similar gains in word retention regardless of task type. Third, this particular design would provide a chance to see which type of input is more enhancing depending on task type.

When more involving tasks (sentence writing and gap filling) were compared with their equivalents within the lower-intermediate group, it was observed that the type of input plays an essential role in learning of form. The results of both productive vocabulary post-tests showed that while the highest amount of gain and retention across the fill-in groups was achieved by the FR group, the SR and SW tasks yielded the greatest productive word gain on the immediate post-test and retention on the delayed post-test respectively. In other words, the superiority of the SR task on the immediate productive vocabulary post-test was overshadowed not only by the SW task but also by the FR task on the delayed productive vocabulary post-test. In a way, this finding concurs with Keating's (2008) study because he found that sentence writing leads to more immediate productive word knowledge, while gap filling enhances active recall of form better over time. He explained the loss in the superiority of sentence writing tasks by stating that such decontextualized tasks might not trigger the kind of processing that is profound enough to retrieve the target forms.

Based on the aforementioned results, it can be concluded that the pairing of the lower involvement task (gap filling) with the written input and the pairing of the higher involvement task (sentence writing) with the audiovisual input are both effective in promoting the retention of form for lower-proficiency learners. There can be several explanations for this finding. First of all, the SW task in this study was performed in a contextualized manner. The learners watched a video podcast and became familiar with the target vocabulary through visual cues. The use of listening as a medium can also account for the higher task achievement because knowledge acquired from listening is more

resistant to decay (Brown, Waring and Donkaewbua, 2008; Hassanzadeh, 2016; van Zeeland and Schmitt, 2013). Mármol and Sánchez-Lafuente (2013) also justified the efficacy of sentence writing in increasing productive word knowledge by proposing that these tasks by nature require employing productive skills rather than receptive ones as benefitted in a reading comprehension task. Thus, the type of knowledge they reinforce inherently can offer an explanation for their success on productive vocabulary post-tests.

On the other hand, the effectiveness of the audiovisual input over the written input in the sentence writing tasks did not apply to the gap filling tasks for the lower-intermediate learners. One reason for this failure might be the poor listening skills of this particular participant group. For most L2 learners, it is often problematic to recognize or segment lexical items especially when learners are exposed to spoken discourse (Vandergrift, 2004; van Zeeland, 2013). Furthermore, when L2 learners have trouble understanding the content, they might often direct their working memory to comprehending the material only rather than focusing on individual words. As a result, their cognitive resources might not be fully available to deal with the unknown words in the discourse (Boers, Eyckmans and Godfroid, 2007). Although Marefat and Hassanzadeh (2014) suggested that audiovisual materials can overcome such drawbacks by offering abundant verbal and pictorial cues, the learners in this study seemed incapable of deploying these clues due to their insufficient linguistic properties and limited text comprehension. What is more, knowledge of meaning is more difficult to be attained through listening (van Zeeland and Schmitt, 2013) and lexical items acquired by means of listening are more likely to be forgotten by less proficient L2 learners (Vidal, 2011). This can explain the significant loss of knowledge that the learners in the FW group exhibited both on the delayed productive and receptive vocabulary post-tests. Considering all these factors, the failure of the FW task in the lower-proficiency group seems to be of no surprise.

With respect to the effects of the written input, it is obvious that it positively affects incidental vocabulary acquisition when it is consolidated with a gap filling task at a lower language level. As Vidal (2011) claimed, reading is a fundamental source of lexical uptake for less proficient L2 learners because speech segmentation in listening tends to be more

complicated for them. Because listening entails real-time processing, learners cannot linger on the words that they do not understand or go back to their first encounter with these words to reflect upon their meanings. However, because the learners in the FR group were provided with a reading passage, they were able to see the target words in their written forms while filling them into the gapped text. This enabled them to dwell on meaning more closely and thus form more durable representations of these words in the long-term memory. Although the FW group was exposed to the target words one more time than the FR group through their video podcast, they could not benefit from this additional exposure due to their poor listening ability. That is why the lower-intermediate learners assigned to the fill-in-by-reading task achieved much higher scores than those assigned to the fill-in-by-watching task.

The results of both receptive vocabulary post-tests pointed out an almost equal performance for the sentence writing groups, a noticeable difference between the gap filling groups and a slight discrepancy across the control groups. When retention of meaning was taken into account, it was found that the SW task resulted in a little more receptive word knowledge (0.5 word) in the long run than the SR task, so it can be concluded that both tasks are equally effective in fostering lexical acquisition regardless of input mode. As Craik and Lockhart (1972) suggested, if L2 learners elaborate on new words more semantically at a deeper level, they tend to recall their meanings more easily than their forms. This accounts for why the learners in both sentence writing groups equally retrieved most of the correct definitions on both receptive vocabulary post-tests while losing a significant amount of their productive word knowledge over time. Therefore, it can be implied that input type does not interfere with vocabulary learning when sentence writing tasks are used with lower-proficiency learners to increase receptive word knowledge. What is more influential is the deep semantic processing that these tasks induce due to their inherent nature, which largely accounts for the retrieval of the target items.

On the other hand, input modality played a distinguishing role between the fill-in tasks in terms of receptive word knowledge. The FR task led to substantially more receptive word gain and retention than the FW task. As mentioned before for productive word knowledge,

the ability of the lower-intermediate learners to recognize individual segments while listening to their video podcast in the FW task seemed to be somewhat limited. Nevertheless, these learners were able to supply more correct meanings than forms although van Zeeland and Schmitt (2013) stated that knowledge of meaning is harder to be acquired through listening. This finding suggested that when the audiovisual input is used with lower-proficiency learners during gap filling tasks, it may not lead to the desired amount of form acquisition partly due to its need for good listening skills. However, it can create a relatively better recall of meaning thanks to the contextual clues it visually presents. This can also explain why the WO group slightly outperformed the RO group although both groups did not tackle with the target words at all. Overall, when both input modes are considered, it can be concluded that the written input seems to be more effective than the audiovisual input in gap filling tasks for less proficient learners.

In summary, it can be concluded that the written input leads to more productive and receptive word knowledge with fill-in tasks in all circumstances, while the audiovisual input leads to more productive and receptive word knowledge with sentence writing tasks over time.

### **5.1.3. Vocabulary Gain from Various Loaded Tasks at Upper-Intermediate Level**

The third research question investigated whether vocabulary tasks with higher levels of involvement load result in more vocabulary gain than tasks with lower levels of involvement load at upper-intermediate proficiency level in terms of productive and receptive word knowledge. To answer this research question, three kinds of vocabulary tasks were created for the upper-intermediate participants. Similar to the vocabulary tasks of the lower-intermediate group, these tasks consisted of a reading/watching-only task (index = 0), a gap filling task (index = 2) and a sentence writing task (index = 3), but they were formed with a different reading passage and a video podcast that were suitable for the level of the high-proficiency learners.

The results of the productive and receptive vocabulary post-tests indicated that sentence writing tasks (index = 3) are more profitable among highly proficient learners than fill-in

tasks (index = 2) which in turn are more effective than reading/watching-only tasks (index = 0) in boosting both productive and receptive word knowledge in the long term. This finding thoroughly verified the predictions of the Involvement Load Hypothesis because Hulstijn and Laufer (2001) also found that the most involving task yields more vocabulary gain and retention than the more involving task which in turn leads to more lexical enhancement than the least involving task.

On the other hand, there is one notable finding worth pointing out for the upper-intermediate learners because this finding contradicted the assumption of the hypothesis concerning the superiority of more involving tasks over less involving ones. In the present study, the fill-in-by-watching (FW) group obtained an equal amount of productive word knowledge as the sentence-writing-by-watching group and an equal amount of receptive word knowledge as the sentence-writing-by-reading group on the relevant immediate post-tests. Compared to the lower-intermediate group, the upper-intermediate learners assigned to the FW task made considerably more initial vocabulary gain owing to their listening ability. Such learners with higher language proficiency possess a larger bulk of lexical knowledge and thus manage to parse the running speech more easily for word-meaning interpretation (Vidal, 2011). This advantage is often unavailable for the learners of lower language proficiency. However, Lin (2010) asserted that having good listening skills may not be satisfactory enough while learning new words as the ability to listen better can only be achieved when it is supplemented with effective reading skills. That is probably why both sentence writing groups outscored the FW group on the delayed post-tests since the learners in the FW group lost a significant number of the target words within two weeks' time. Interestingly, this finding is in contradiction with Vidal's (2011) study. She found that words are less vulnerable to decay when they are acquired through listening at high proficiency levels. Therefore, the efficacy of the FW task at the higher-proficiency level can be attributable to the use of the audiovisual input.

For the higher-proficiency group, the results entirely validated the hypothesis by stressing the superiority of sentence writing tasks over gap filling tasks followed by reading/watching-only tasks in increasing both productive and receptive word knowledge.

In general, the findings associated with the higher-proficiency group in the present study concurred with former studies conducted with intermediate or advanced L2 learners (Hulstijn and Laufer, 2001; Kim, 2011) and emphasized the strength of the evaluation component in task layout. Although Hulstijn and Laufer (2001) regarded all the components to be identical in value, Kim (2011) noted that strong evaluation may have a much greater impact on the acquisition of L2 vocabulary. That is why the sentence writing groups in this study outperformed all the other task groups at upper-intermediate level because their tasks induced stronger evaluation. These learners were more cognitively challenged while they were combining the target words with others to form original sentences during their tasks. In doing so, they were able to commit these target items to memory better since they were more involved in the task. This shows that sentence writing tasks are also effective at a higher L2 proficiency level.

Another reason for the efficacy of sentence writing tasks at upper-intermediate level might stem from the link between L2 proficiency and cognitive awareness. Research has indicated that learners of high L2 proficiency tend to develop and employ more cognitive strategies since linguistic knowledge is the key factor for these strategies to emerge and operate effectively (Baker, 2005; Griffith and Ruan, 2005). When L2 learners have a better command of the language, they can ease their mental processing and direct their remaining cognitive resources to higher-order skills. That is why the higher-proficiency group gained the most vocabulary knowledge through the sentence writing tasks although these tasks induced more cognitive load with strong evaluation. Thanks to their grammatical competence and L2 vocabulary size, these learners could overcome the cognitive demand and linguistic difficulty of their task.

On the other hand, when less proficient L2 learners are considered, more cognitive resources are needed for task completion because these learners lack certain language skills, have limited processing capacity and cannot decode words automatically (Boers, Eyckmans and Godfroid, 2007; de Bot, 1992; Vandergrift, 2004). Due to their insufficient L2 competence, they also make use of fewer cognitive strategies and employ them less successfully (Anderson, 1991; Green and Oxford, 1995). That might explain why the gap

filling task (by reading) was as effective as the sentence writing task among the lower-proficiency learners. Lu (2013) also acknowledged that lower-intermediate learners can benefit more from blank-filling tasks because highly loaded tasks might take more time to complete making them inconvenient for EFL classrooms due to time constraints.

As mentioned before, the lowest vocabulary gain was obtained by the control groups at upper-intermediate level. Still, their reading/watching-only tasks resulted in more receptive word knowledge than productive word knowledge. This suggests that without any explicit focus on the target words during task implementation, it seemed rather difficult for these learners to acquire the form. However, meaning was still acquired to some extent even though these learners were not provided with any enhancing or noticing effects that helped them divert their attention to the target items. Thus, it can be concluded that indirect exposure to unfamiliar words can somewhat lead to their acquisition to a certain degree if learners are willing to employ their inference or guessing strategies as in the case of the learners in the control groups who were used to doing such strategy-based practices in their language classrooms.

In summary, the results of the upper-intermediate learners fully validated the predictions of the Involvement Load Hypothesis since the more involving task (sentence writing) yielded greater vocabulary retention than the less involving task (gap filling) which in turn led to more vocabulary gain than the least involving task (reading/watching-only).

#### **5.1.4. Vocabulary Gain from Different Modes of Input at Upper-Intermediate Level**

The final research question explored whether vocabulary tasks with different input modalities (audiovisual versus written) yield differences in vocabulary gain at upper-intermediate proficiency level in terms of productive and receptive word knowledge. To address this research question, three pairs of tasks that shared an identical index of involvement load were designed either with the written input (a reading passage) or the audiovisual input (a video podcast) to be administered to the upper-intermediate learners.

The results of the immediate productive vocabulary post-test showed that the audiovisual input resulted in more initial vocabulary gain within the sentence writing and gap filling

groups, while no input difference was found between the control groups. Two weeks later, the results pertaining the sentence writing and control groups remained the same, whereas the written input led to slightly more vocabulary retention with the fill-in task. As for the results of the immediate receptive vocabulary post-test, it was observed that the audiovisual input was more effective than the textual input regardless of the task type. When the retention of meaning over time was considered, it was seen that the superiority of the video input continued to exist within the gap filling and control groups, while both types of input were equal in terms of their contribution to receptive word retention with the sentence writing tasks.

In general, the findings of the higher-proficiency group suggested that the audiovisual input seems to yield more vocabulary gain and retention than the written input. These findings lend empirical support to those of Marefat and Hassanzadeh's (2014) and Hassanzadeh's (2016) studies. They also found that L2 learners are able to increase their lexical knowledge more through vocabulary tasks with the video input than comprehension-based reading or listening activities. As Marefat and Hassanzadeh (2014) suggested, the written input can be overwhelming sometimes, so when L2 learners are presented with a reading text all at once, they may simply disregard unfamiliar words for the sake of the gist. With the oral input, they might fail to recognize the useful paralinguistic features in the speech discourse. However, vocabulary achievement can be higher with audiovisual aids because they can facilitate lexical acquisition by providing plentiful contextual clues and thus establishing deeper mnemonic bonds than any other medium. This suggests that the effect of the audiovisual input on lexical uptake seems to be stronger than the effect of other input types.

Based on the nature of the audiovisual input, several more explanations can be made to account for its superiority. Vandergrift (2004) asserted that the video input provides L2 learners with a great number of pictorial aids that help activate their schemata and build up their background knowledge about the relevant theme (i.e., top-down processing). Through the presentation of images, learners can contextualize the information they are attending to. This can help reduce the cognitive load on their attentional mechanisms and divert their mental resources to other features of spoken discourse such as segmenting the speech (i.e.,

bottom-up processing). Thus, learners can comprehend the spoken text better and infer the meaning of unknown vocabulary in the input. However, the visual clues in the audiovisual input are not present in the written input. That is why most of the upper-intermediate learners who were assigned to read a passage acquired fewer target words than those who were instructed to watch a video podcast. This was the case with the lower-proficiency participants too when the video input was paired with the sentence writing task. However, the textual input enhanced more vocabulary gain when it was presented with a gap filling task at lower-intermediate level. This implies that L2 proficiency is an influential factor in determining how different kinds of input affect task effectiveness and incidental vocabulary acquisition.

One remarkable common finding concerning both levels of language proficiency is the efficacy of the sentence writing task when it is implemented with the audiovisual input. Based on Swain's (1985) Output Hypothesis, writing individual sentences with the target words can be regarded as an output task because it requires learners to produce the language. As mentioned before, the audiovisual input fosters the learning of unfamiliar lexical items more than the written input. Therefore, when this input type is paired with an output task, it facilitates incidental vocabulary acquisition even more because sentence writing provides learners with the opportunity to experiment with the target words that they have just heard from a video podcast. While forming sentences that somehow indicate the definitions of the target items, learners can compensate for the gaps in their lexicon by concentrating on meaning. This helps them process the target words more deeply, master them better and automatically retrieve them when they are needed to convey a particular message during communication. In this way, sentence writing tasks reinforce the fluent use of newly met words. As clearly shown in this study, they can be even more efficient when they are performed through an audiovisual resource that supplies learners with abundant visual cues to understand the context, usage and meaning of each target word.

On the other hand, a notable distinction emerged between the two proficiency groups when the efficacy of the gap filling task was compared based on its input presentation. While the fill-in-by-watching (FW) task was one of the most effective tasks among the higher-

proficiency learners, it was one of the least successful tasks within the lower-intermediate group. This difference might stem from the way input modality acts when it is interacted with L2 proficiency. In the FW task, the learners were instructed to fill in the blanks left out in the sentences extracted from the video podcast while viewing it. The video podcast offered a number of visual cues, but the task required the learners to listen carefully in order to identify the target words in the speech stream. As spoken discourse involves continuous speech (i.e., no clear-cut boundaries between the words), it necessitates instant processing. This implies that listeners basically have less time to comprehend the input than readers who usually have control over their written texts. In addition, speech segmentation conveys a particular difficulty for L2 learners since they often miss the upcoming or forget the preceding parts of the listening text (Goh, 2000). Thus, they often try to listen for the main idea, which makes it hard to notice the newly met words in the input. When a low level of language proficiency is added to the picture, problems might occur with the acquisition of unfamiliar lexical items (Vidal, 2011). That is why the more competent learners in this study who completed their FW task attained more vocabulary gain and retention than the less proficient learners who performed the same type of task but struggled harder due to their poorer listening skills.

In summary, although Hulstijn (2003) asserted that “it is elaboration of or involvement in the lexical information being processed rather than any of the factors per se that determines retention” (p. 364), empirical evidence from the present study suggests that input type is an important variable that influences incidental vocabulary acquisition. Each task in this study was planned with a corresponding task that shared an identical index of involvement load but comprised a different type of input mode. Based on Laufer and Hulstijn’s (2001) claim that tasks constituted with the same level of involvement load are equally effective in enhancing L2 word retention, the tasks in this study were supposed to result in similar word gains. However, it was revealed that the tasks created with the audiovisual input yielded more vocabulary gain among the higher-proficiency group. Similarly, input modality was influential in distinguishing between the tasks that involved an equal degree of involvement load among the lower-intermediate learners. For that participant group, it was concluded

that the gap filling task with the textual input seemed to be more efficient, whereas the sentence writing task with the video input was more effective. All these findings contradict with Laufer and Hulstijn's (2001) prediction on identically loaded tasks leading to similar amounts of lexical uptake because it is evident that task efficacy is influenced by input type which in turn is affected by L2 proficiency. Therefore, it can be suggested that input modality might be treated as a fourth component to the Involvement Load Hypothesis when task design is considered.

## **5.2. Summary of the Findings and Discussion**

The present study attempted to shed light on the effects of task-induced involvement load, input type and language proficiency on incidental vocabulary acquisition by means of six vocabulary tasks that operationalised three degrees of involvement load and two different modes of input. The learners who completed these tasks were compared on their immediate and delayed productive and receptive post-test scores at two proficiency levels.

First of all, fully concurring with the predictions of the Involvement Load Hypothesis, the results favoured the tasks with the highest level of involvement load (sentence writing) over the tasks that are less involving (gap filling) or induce no learner involvement (reading alone or video comprehension). Concerning this finding, it can be inferred that strong evaluation has a much stronger impact on the initial gain and retention of new words. In general, the findings imply that the construct of task-induced involvement load can be beneficial while predicting the efficacy of particular tasks over others.

Secondly, the participants acquired varying amounts of vocabulary gain and retention from the tasks that had an identical index of involvement load. This finding contrasts with Laufer and Hulstijn's (2001) claim for equally loaded tasks that lead to similar vocabulary knowledge. Since the contradictory finding of the present study applies to both levels of language proficiency, the different word gains of the learners who performed identically loaded tasks can be attributed to the influence of input modality on incidental lexical acquisition.

Thirdly, the present study highlighted the distinctive role that input type plays in enhancing lexical uptake when task type and L2 proficiency intervene in the learning process together. The results of the lower-intermediate participants showed that the written input is more effective when it is combined with a gap filling task, while the audiovisual input facilitates the acquisition of the target words when it is consolidated with a sentence writing task. This finding is explained by the learners' limited language abilities that might remain insufficient to help interpret the contextual clues in audiovisual materials. As for the upper-intermediate participants, the findings pointed out the superiority of the audiovisual input. The learners who performed their assigned task by viewing a video podcast achieved greater scores than those who completed their task by reading a passage. These findings suggest that audiovisual resources can be an effective way of teaching unfamiliar words in language learning contexts.

In conclusion, the current study provided empirical support for the assumptions of the Involvement Load Hypothesis. It primarily indicated that task efficacy is positively affected when vocabulary tasks induce higher levels of learner involvement and originate from an audiovisual source. Based on these noteworthy findings, several important pedagogical implications can be made for teachers and material designers.

### **5.3. Pedagogical Implications**

The current study attempted to test the predictions of the Involvement Load Hypothesis by showing how vocabulary tasks can enhance incidental vocabulary acquisition. Therefore, the findings can offer some fundamental implications for teachers to improve their classroom dynamics and for material developers to design more efficient instructional resources that facilitate L2 vocabulary acquisition.

While trying to confirm the assumptions of the hypothesis, the present study emphasized the superiority of word-focused vocabulary tasks in promoting lexical uptake over mere reading or video comprehension activities. As the sentence writing and gap filling tasks fostered more word retention than the reading/watching-only tasks in the study, it can be concluded that explicit focus is crucial to the acquisition of unknown words (Laufer, 2003).

Since the participants of the study were not told to memorize the target words or forewarned about the upcoming vocabulary post-tests, it is obvious that they acquired most of the target items as a result of getting involved in their vocabulary task. Engaging L2 learners in such tasks that instruct them to use the newly encountered words can make them notice these lexical items and thus yield more vocabulary gain and retention. Previous research also draws attention to the greater benefits of using word-focused vocabulary tasks over reading alone in leading to more lexical knowledge (Hulstijn and Laufer, 2001; Knight, 1994; Laufer, 2003; Llach, 2009; Min, 2008; Paribakht and Wesche, 1997). Therefore, rather than relying only on reading (or listening) comprehension activities to teach new vocabulary, teachers should try to make use of additional vocabulary tasks as often as possible. As for material developers, they should be more encouraged to devise such tasks to be included in their language coursebooks.

Secondly, although Hulstijn and Laufer (2001) regarded all the three components of the hypothesis (need, search and evaluation) to be equal in weight, it is revealed that strong evaluation might have a much stronger impact on the total index of a task, as Kim (2011) proposed. Thus, designing tasks with higher degrees of evaluation can enable L2 learners to acquire more words and recall them later with less mental effort. However, it is critical to note here that some tasks that demand higher learner involvement (e.g., composition writing) also require deeper cognitive processing, a bigger L2 vocabulary size and more complex linguistic abilities, so extreme caution should be taken when implementing such tasks with learners of limited language competence. Although Keating (2008) suggested that highly loaded tasks can be employed to low-proficiency learners too, the acquisition and retention rates of unfamiliar words will be low if learners are not able to meet the task requirements because of their inadequate lexicon or grammatical competence. As a result, teachers should be careful while choosing appropriate vocabulary tasks that are not cognitively and linguistically too demanding or overwhelming for their students. The inefficacy of the fill-in-by-watching task for the lower-intermediate learners in this study provides strong support for this warning as the learners could not perform the task the way they were asked to due to their inadequate listening skills. Therefore, as Hulstijn and Laufer

(2001) implied, it might be wiser to teach new words to less competent L2 learners with moderate evaluation rather than strong evaluation.

A third implication of the present study relates to the impact of input type on incidental vocabulary acquisition along with the issue of identically loaded tasks. As the results indicated, tasks that share an identical load of involvement do not result in similar amounts of vocabulary gain and retention despite Laufer and Hulstijn's (2001) claim. Thus, this contradictory finding is attributable to the implementation of the same tasks through different varieties of input. As it is clearly seen in this study, the amount of lexical uptake varies depending on both the kind of input that is presented to the learners and the language proficiency that these learners possess. It is concluded that the audiovisual input seems to yield better results in incidental vocabulary acquisition among highly proficient learners. However, its effects are visible among learners of low proficiency only when it is paired with an output task that involves strong evaluation (e.g., sentence writing). All in all, these findings suggest that video podcasts can be a useful tool in teaching L2 vocabulary, but teachers tend to avoid using such audiovisual aids because of time constraints, lack of technological equipment, topic difficulty or irrelevancy and so on. However, authentic materials such as video podcasts offer a really motivating and engaging content for L2 learners as well as being an instructional source for teaching unknown words. Therefore, teachers should make an effort to exploit such beneficial resources in their classrooms.

With regard to the written input, it seems to promote more lexical acquisition for less proficient L2 learners when it is blended with a fill-in task. This conclusion should encourage teachers to use reading texts in their classrooms in order to create a contextualized environment for learning new words. As many teachers know, language coursebooks often consist of vocabulary exercises that require learners to supply the words given in a box into gapped sentences provided below that box. However, this kind of vocabulary learning lacks a meaningful context because most of these sentences are decontextualized or irrelevant. Nevertheless, fill-in-the-blank exercises are preferred as one of the most common types of vocabulary activities both in coursebooks and examinations since they save time and assess vocabulary knowledge easily. However, Keating (2008)

claimed that decontextualized tasks might stimulate the kind of processing that is likely to be too shallow for the retention of the unknown words. Therefore, gap filling tasks complemented with reading comprehension can be more effective in the retrieval of newly acquired words.

For material designers, there are several implications that can be put into practice. First of all, material developers can prepare vocabulary tasks that involve different degrees of task-induced involvement load and various types of input to ensure that different learners can benefit differently from these resources depending on their language skills, cognitive abilities and instructional needs. For example, for less competent L2 learners, less demanding vocabulary tasks that include gap filling supplemented with reading comprehension can be recommended, while a video viewing activity accompanied by sentence or composition writing can be designed for more advanced L2 learners. Moreover, material designers can supply learners with numerous reading texts and rich audiovisual resources that are constituted from their coursebooks. This will be really advantageous especially for teachers because they often have a hard time finding the most suitable materials that comply with their language books or meet their learners' needs in terms of theme, organization or language. Therefore, such teachers will not lose time trying to adapt their supplementary resources to their goals.

Finally, the present study offers valuable implications for teacher education. Its results can enable prospective language teachers to understand the value of word-focused tasks and gain a better insight into task design before they start devising their own vocabulary tasks. Similarly, teacher educators can refer to the findings of the present study while explaining their students why they should choose particular tasks but not others for vocabulary teaching. Thus, teacher candidates will feel self-sufficient in the future when a need for such instructional activities emerges not only for in-class practices but also for institutional examinations.

To summarize, language teachers, material developers and teacher educators should consider the effects of the construct of task-induced involvement load, input type and L2

proficiency on incidental vocabulary acquisition while preparing or choosing instructional vocabulary tasks for L2 learners. Creating tasks that best suit learners' needs, language competence and cognitive capacity is extremely important. Otherwise, implementing cognitively or linguistically challenging tasks to learners of poor language skills will be futile since learners might not acquire any new words. Therefore, it is essential to take all these factors into account in order to devise valuable vocabulary tasks that can foster more L2 lexical knowledge.

#### **5.4. Limitations and Suggestions for Further Research**

The present study has several limitations. First of all, the number of learners that participated in this study is 236. Further research can be conducted with a relatively larger size of participants at various age groups and proficiency levels in order to verify the findings regarding the Involvement Load Hypothesis.

Secondly, the current study employed only three major kinds of vocabulary tasks with two types of input. The findings deduced from this comparison can offer illuminating answers for similar studies, but they cannot make valid generalizations that fully confirm the assumptions of the hypothesis. Therefore, future studies should set out to examine the effects of many other vocabulary tasks designed with different/equal degrees of task-induced involvement load and various modes of input on incidental vocabulary acquisition. For example, a vocabulary task that requires learners to write novel sentences by incorporating the target words presented through a video can be compared with the one that instructs learners to orally produce original sentences by using the target words provided in the same video. These two tasks would include an equal level of involvement load and the same kind of input presentation, but they would impose the use of different productive language skills (writing and speaking).

Additionally, the current study did not investigate the search component of the hypothesis as the tasks would be too demanding for the participants if they had to search for the definitions of the target words in a bilingual dictionary. Then it would be necessary to teach dictionary look-up skills to all the learners since many of them do not know how to use a

dictionary properly. However, future research can look at the effects of search on L2 lexical acquisition by making different combinations of the three components as well as integrating the written, oral and/or audiovisual input.

A fourth important limitation of the present study lies in the selection of the target words. The target items were chosen from their relevant video podcasts that were originally supplied in the supplementary materials of the learners' coursebooks. The selection was carried out in this way because if the video podcasts were taken from an external resource, the content and language might be too confusing for the learners. The limitation concerning the target words stems from their orthographic forms since some of them are alike in spelling (e.g., *pager* and *patrol* for the lower-intermediate learners and *complacent* and *conscientious* for the upper-intermediate learners). Moreover, it was observed from the results of the productive vocabulary post-tests that a few of the words (e.g., *treacherous*, *conscientious*, *abundance* and *dedicated*) were particularly difficult for the learners to recall their written forms. However, L2 learners sometimes have to learn and produce such words that are long in form. Besides, in order to eliminate the factor of word length, the number of long words was balanced in both proficiency groups. Moreover, the learners' partial answers on the productive vocabulary post-tests were also evaluated. Thus, it can be concluded that the inclusion of orthographically complex words did not necessarily manipulate the findings. However, the number of target words that appeared during each task could have been kept constant to avoid any interfering effect of exposure.

Similarly, word category was balanced across proficiency in order to prevent its effects on the acquisition of the target items. The target words selected for both levels of language competence consisted of three nouns, three adjectives and two verbs. Balancing word type in both groups enabled a reliable assessment. However, further research can study the effects of word type on task effectiveness and incidental vocabulary acquisition by including only nouns, verbs or adjectives.

Another confining limitation of the study is related to the factor of time-on-task (i.e., the time invested by learners to finish a task). Tu (2004) stated that tasks with higher learner

involvement may be more efficient not because they induce a deeper level of processing, but because learners are exposed to the target words much longer as a result of spending more time during their task. Thus, to balance the variable of time-on-task in the present study, several comprehension questions were added to certain tasks. Nevertheless, in order for the tasks to be performed properly, varying amounts of time were allocated for each task because of task nature. This does not pose a threat for the interpretation of the results since Hulstijn and Laufer (2001) considered the element of time-on-task to be an inherent part of a task. It is not regarded as a separate factor. Moreover, it is not logical to assume that the completion of a sentence writing task would last the same as doing a reading comprehension task. Therefore, the present study did not keep the time-on-task variable under control, but future studies might investigate the efficacy of involving tasks with the time controlled versus uncontrolled.

Finally, the present study only administered two types of vocabulary post-tests to measure the participants' productive and receptive vocabulary gain and retention. Future studies in the field can employ different kinds of recognition and recall tests to assess both productive and receptive word knowledge acquired through the construct of task-induced involvement load. For example, a word-form recognition test (circling the target words presented in a different text) or a meaning-translation test (translating the sentences that incorporate the target words) can be designed depending on the type of knowledge to be measured.

## REFERENCES

- Akbulut, Y. (2007). Variables predicting foreign language reading comprehension and vocabulary acquisition in a linear hypermedia environment. *The Turkish Online Journal of Educational Technology*, 6(1), 53-60.
- Albrechtsen, D., Haastrup, K., & Henriksen, B. (2008). *Vocabulary and writing in a first and second language: Processes and development*. Basingstoke: Palgrave Macmillan.
- Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal*, 75, 460-472.
- Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (pp. 89-195). New York: Academic Press.
- Baddeley, A. D. (1978). The trouble with levels: A re-examination of Craik and Lockhart's framework for memory research. *Psychological Review*, 85(3), 139-152.
- Baddeley, A. D. (1999). *Essentials of human memory*. Hove: Psychology Press.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47-89). New York: Academic Press.
- Baker, L. (2005). Developmental differences in metacognitive knowledge: Implications for metacognitively oriented reading instruction. In E. Israel, C. Block, L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development* (pp. 61-79). London: Routledge.
- Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, 52, 323-363.
- Boers, F., Eyckmans, J., & Godfroid, A. (2007). Catering for limited processing capacity to foster incidental vocabulary uptake. In K. Pelsmaekers & C. Rollo (Eds.), *Economically speaking: Essays in honour of Chris Braecke* (pp. 169-187). Antwerp: Garant Publishers.
- Bonk, W. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14, 14-31.
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20, 136-163.

- Coady, J., & Huckin, T. (1997). *Second language vocabulary acquisition*. Cambridge: Cambridge University Press.
- Craik, F., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Craik, F., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268-294.
- De Bot, K. (1992). Applied linguistics. *Applied Linguistics*, 13, 1-24.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1999). Factors in the incidental acquisition of second language vocabulary from oral input. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 35-62). Amsterdam: John Benjamins.
- Eysenck, M. W. (1978). Levels of processing: A critique. *British Journal of Psychology*, 69, 157-169.
- Fatalaki, J. A. (2014). Involvement load hypothesis: Word meaning retention across oral and written task types. *International Letters of Social and Humanistic Sciences Online*, 37, 29-45.
- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40(2), 273-293.
- Gass, S. (1999). Incidental vocabulary learning. *Studies in Second Language Acquisition*, 21(2), 319-333.
- Gernsbacher, M., Varner, K., & Faust, M. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 430-445.
- Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28, 55-75.
- Goulden, R., Nation, P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11(4), 341-363.
- Grabe, W., & Stoller, F. (2002). *Teaching and researching reading*. Harlow, UK: Longman.
- Green, J. M., & Oxford, R. L. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Quarterly*, 29, 261-297.

- Griffith, P. L., & Ruan, J. (2005). What is metacognitive knowledge and what should be its role in literacy instruction. In E. Israel, C. Block, L. Bauserman, & K. Kinnucan-Welsch (Eds.), *Metacognition in literacy learning: Theory, assessment, instruction, and professional development* (pp. 3-18). London: Routledge.
- Hassanzadeh, M. (2016). Experimenting the influence of input modality on involvement load hypothesis. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 10(6), 2043-2048.
- Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English texts. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary acquisition* (pp. 130-152). Norwood, NJ: Ablex.
- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21, 303-317.
- Henriksen, B., Albrechtsen, D., & Haastrup, K. (2004). The relationship between vocabulary size and reading comprehension in the L2. In D. Albrechtsen, K. Haastrup, & B. Henriksen (Eds.), *Writing and vocabulary in foreign language acquisition* (pp. 129-140). Copenhagen: Museum Tusculanum Press.
- Hincks, R. (2003). Pronouncing the academic word list: Features of L2 student oral presentations. In *Proceedings of the 15<sup>th</sup> International Congress of Phonetics Sciences* (pp. 1545-1548). Barcelona: Universitat Autònoma de Barcelona.
- Hinkel, E. (2001). Giving personal examples and telling stories in academic essays. *Issues in Applied Linguistics*, 12, 149-170.
- Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8, 689-696.
- Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language*, 11(2), 207-223.
- Hu, H. P., & Deng, L. J. (2007). Vocabulary acquisition in multimedia environment. *US-China Foreign Language*, 5(8), 55-59.
- Hu, M., & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403-430.
- Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition*, 21(2), 181-193.

- Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 349-381). Oxford: Blackwell.
- Hulstijn, J., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal*, 80(3), 327-339.
- Hulstijn, J., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539-558.
- Jing, L., & Jianbin, H. (2009). An empirical study of the involvement load hypothesis in incidental vocabulary acquisition in EFL listening. *Polyglossia*, 16, 1-11.
- Joe, A. (1998). What effect do text-based tasks promoting generation have on incidental vocabulary acquisition? *Applied Linguistics*, 19(3), 357-377.
- Jones, L. (2004). Testing L2 vocabulary recognition and recall using pictorial and written test items. *Language Learning and Technology*, 8(3), 122-143.
- Karalık, T. (2016). *The effect of task-induced involvement load on incidental vocabulary acquisition among ELT students* (Unpublished master's thesis). Anadolu University, Turkey.
- Keating, G. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, 12(3), 365-386.
- Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, 61(1), 100-140.
- Kim, D., & Gilman, D. A. (2008). Effects of text, audio, and graphic aids in multimedia instruction for vocabulary learning. *Educational Technology & Society*, 11(3), 114-126.
- Kintsch, W., & Kozminsky, E. (1977). Summarizing stories after reading and listening. *Journal of Educational Psychology*, 69(5), 491-499.
- Knight, S. M. (1994). Dictionary use while reading: The effects on comprehension and vocabulary acquisition for students of different verbal abilities. *The Modern Language Journal*, 78(3), 285-299.

- Krashen, S. D. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, 73, 440-464.
- Krashen, S. D. (1993). The case for free voluntary reading. *The Canadian Modern Language Review*, 50(1), 72-82.
- Kweon, S. O., & Kim, H. R. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language*, 20(2), 191-215.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan.
- Laufer, B. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 140-155). Cambridge: Cambridge University Press.
- Laufer, B. (2003). Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence. *Canadian Modern Language Review*, 59(4), 567-587.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22, 1-26.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Lin, L. F. (2010). English learners' incidental vocabulary acquisition in the video-based CALL program. *The Asian EFL Journal*, 12(4), 51-66.
- Llach, M. P. A. (2009). The effect of reading only, reading and comprehension, and sentence writing in lexical learning in a foreign language: Some preliminary results. *RESLA*, 22, 9-33.
- Lockhart, R. S., & Craik, F. I. M. (1978). Levels of processing: A reply to Eysenck. *British Journal of Psychology*, 69, 171-175.
- Lu, M. (2013). Effects of four vocabulary exercises on facilitating learning vocabulary meaning, form, and use. *TESOL Quarterly*, 47, 167-176.
- Maleki, N. A. (2012). The effect of the involvement load hypothesis on improving Iranian EFL learners' incidental vocabulary acquisition in listening comprehension classes. *Australian Journal of Basic and Applied Sciences*, 6(9), 119-128.
- Marefat, F., & Hassanzadeh, M. (2014). Vodcast: A breakthrough in developing incidental vocabulary learning. *Iranian Journal of Applied Linguistics*, 17(2), 27-58.

- Mármol, G. A., & Sánchez-Lafuente, A. A. (2013). The involvement load hypothesis: Its effect on vocabulary learning in primary education. *RESLA*, 26, 11-24.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-97). Bristol: Multilingual Matters.
- Min, H.-T. (2008). EFL vocabulary acquisition and retention: Reading plus vocabulary enhancement activities and narrow reading. *Language Learning*, 58, 73-115.
- Nagy, W. (1997). On the role of context in first and second language vocabulary learning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 64-83). Cambridge: Cambridge University Press.
- Nagy, W., & Herman, P. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. McKeown & M. Curtis. (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Erlbaum.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Heinle & Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review*, 63(1), 59-81.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). New York, NY: Cambridge University Press.
- Nation, I. S. P., & Coady, J. (1988). Vocabulary and reading. In R. Carter & M. McCarthy (Eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.
- Nation, I. S. P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12(2), 355-380.
- Nelson, T. O. (1977). Repetition and depth of processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 151-171.
- Neuman, B., & Koskinen P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly*, 27(1), 96-106.

- Norman, D. A. (1976). *Memory and attention: an introduction to human information processing*. London: John Wiley.
- Nunan, D. (1999). *Second language teaching and learning*. Boston: Heinle & Heinle.
- Paribakht, T. S., & Wesche, M. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11(1), 9-29.
- Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp.174-200). New York: Cambridge University Press.
- Paribakht, T. S., & Wesche, M. (1999). Reading and 'incidental' L2 vocabulary acquisition. *Studies in Second Language Acquisition*, 21, 195-229.
- Peters, E., Hulstijn, J., Sercu, L., & Lutjeharms, M. (2009). Learning L2 German vocabulary through reading: The effect of three enhancement techniques compared. *Language Learning*, 59, 113-151.
- Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18(1), 1-28.
- Rodgers, M. (2013) *English language learning through viewing television: An investigation of comprehension, incidental vocabulary acquisition, lexical coverage, attitudes, and captions* (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129-158.
- Schmidt, R. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.
- Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal*, 36(2), 139-152.
- Sternberg, R. J. (1987). Most vocabulary is learnt from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-106). Hillsdale, NJ: Erlbaum.

- Sydorenko, T. (2010). Modality of input and vocabulary acquisition. *Language Learning & Technology*, 14(2), 50-73.
- Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-256). New York: Newbury House.
- Tannenbaum, K. R., Torgesen J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10(4), 381-398.
- Teng, F. (2015). EFL vocabulary learning through reading BBC news: An analysis based on the Involvement Load Hypothesis. *English as a Global Language Education (EaGLE) Journal*, 1(2), 63-90.
- Tu, H. F., & Su, I. R. (2004). *Effects of task-induced involvement on incidental vocabulary learning in a second language*. Unpublished Master Thesis. National Tsing Hua University, Taiwan.
- Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381-403). New York: Academic Press.
- Tulving, E., & Patterson, R. D. (1968). Functional units and retrieval processes in free recall. *Journal of Experimental Psychology*, 77(2), 239-248.
- Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics*, 24, 3-25.
- VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition*, 12, 287-301.
- van Zeeland, H. (2013). Vocabulary and listening. In C. A. Chapelle (Ed.), *The Encyclopaedia of Applied Linguistics* (pp. 1-6). Australia: Wiley-Blackwell.
- van Zeeland, H., & Schmitt, N. (2012). Lexical coverage and L1 and L2 listening comprehension: The same or different from reading comprehension. *Applied Linguistics*, 34(4), 457-479.
- van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening. *System*, 41, 609-624.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219-258.
- Walker, L. (1976). Comprehending writing and spontaneous speech. *Reading Research Quarterly*, 11, 144-167.

- Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15(2), 130-163.
- Webb, S., & Rodgers, M. P. H. (2009). The lexical coverage of movies. *Applied Linguistics*, 30(4), 407-427.
- Wesche, M. B., & Paribakht, T. S. (2000). Reading-based exercises in second language vocabulary learning: An introspective study. *The Modern Language Journal*, 84(2), 196-213.
- Wodinsky, M., & Nation, P. (1988). Learning from graded readers. *Reading in a Foreign Language*, 5(1), 155-161.
- Yaqubi, B., Rayati, R. A., & Gorgi, N. A. (2010). The involvement load hypothesis and vocabulary learning: The effect of task types and involvement index on L2 vocabulary acquisition. *The Journal of Teaching Language Skills (JTLS)*, 2(1), 145-163.
- Zahar, R., Cobb, T., & Spada, N. (2001). Acquiring vocabulary through reading: Effects of frequency and contextual richness. *The Canadian Modern Language Review*, 57, 541-572.
- Zimmerman, C. (1997). Do reading and interactive vocabulary instruction make a difference? An empirical study. *TESOL Quarterly*, 31, 121-140.



## **APPENDICES**

## APPENDIX A

### READING-ONLY (RO) TASKS

#### LOWER-INTERMEDIATE LEVEL:

##### BEFORE READING

##### A. Discuss the question below in class.

The sea can be a dangerous place. What risks are there for...?

- surfers and swimmers
- fishermen and sailors

##### WHILE READING

##### The Working Week

##### *Free time by the sea*

Cornwall is in the south-west of England. It is famous for its stunning scenery, pretty villages and beautiful beaches. And there's plenty to do with your free time here. For some people, the weekend is about taking the time to relax. At the beach, people take their dogs for walks, catch up with friends or spend time with their families. And for others, it's all about getting active. You can be a successful accountant during the week and a passionate surfer at the weekend. You can go for a swim or perhaps explore the coastline by kayak. But while the sea and coastline are beautiful, they can also be treacherous, with strong currents and unpredictable weather conditions.

##### *The RNLI*

Every year, thousands of people get into difficulty on the coast, and it's not just the holiday-makers. In the UK, when something goes wrong at sea or on the beach, there's an amazing group of dedicated people who give up their free time to help: the members of the Royal National Lifeboat Institution, the RNLI. It is a charity that saves lives at sea. The RNLI was formed in 1824 to help anyone who got into trouble in the seas around the coastline of the UK. It has both lifeboat volunteers who respond to emergency calls, and lifeguards who patrol more than 160 beaches. The charity uses money donated by the public to pay for its work.

Since it started, the RNLI's boats and lifeguards have saved more than 139,000 lives. Today, there are over four and a half thousand lifeboat crew and over 900 seasonal lifeguards, working for the organization. Most of these people are volunteers who put their own lives in danger to prevent accidents and rescue people in need. For RNLI lifeguards, 95 per cent of the job is to prevent accidents before they happen at the beach. They watch the sea carefully to check that no one is in trouble. They only stay on duty for short periods of time before they have to take a break. Lifeboat crew members attend emergencies at sea. Their pager will go off, and they will rush down to the lifeboat station and launch the lifeboat.

##### *RNLI volunteers*

What kind of people volunteer for the RNLI? Robin Howell is a senior lifeguard on Perranporth beach in Cornwall. It's 10 a.m. on a Sunday in June, and Robin is starting

work. He's setting up all his equipment and getting ready for a day keeping people safe on the beach. He has a radio to stay in contact with the lifeguard base. Robin has another job. He is a geography teacher at a secondary school in Cornwall. He works all year round as a teacher. As a teacher, he doesn't work at the weekends, but from March all the way through to September he works as a lifeguard at the weekends. His lifeguard hours are between 10 and 6 p.m. in the evening.

Having two jobs doesn't leave much time for relaxing, and even when he's not working, Robin likes to stay busy. He says, "I think the main problem with having two jobs is just literally trying to fit it in the space of a week, but I'd rather be busy than sitting at home watching TV. There's not a lot of free time, but I try to do as much sport as possible. I swim in the mornings at 6.30. I have two jobs because I love both of them. I love the lifeguarding because it's outside, it's very physical and I love the teaching because I just like the interaction with the pupils and the fact that I can try and make a difference to their education."

25 year-old Ben Eglington is part of the RNLI lifeboat crew in the town of Newquay in Cornwall. His crew is made up of many volunteers from the local community. There are fishermen, builders, landscape gardeners, hotel/pub owners and fishmongers. Ben's been part of the crew for eight years. He's a busy man, who works full-time as an engineer in the Royal Navy during the week. During the week, he works away in Portsmouth, which is 210 miles away. Every weekend, he makes the epic journey back to Newquay to help out with the RNLI, but Ben doesn't mind giving up his free time to help. He says, "I volunteer my time with the lifeboat because it's good to give something back to the community. It's also a great achievement when you're out there and you actually save somebody's life."

It's thanks to Robin, Ben and all the amazing RNLI volunteers that the beautiful beaches, coastline and seas around the UK are safer for everyone to enjoy.

**B. Read *Free time by the sea*. Make notes to answer the questions.**

- 1) What relaxing activities can you do in Cornwall?
- 2) Name three active sports you can do there.
- 3) What are some of the dangers of the Cornish coast?

**C. Read *The RNLI*. Match the numbers to what they refer to.**

- |                      |                                                     |
|----------------------|-----------------------------------------------------|
| 1) 160               | a) the number of people whose lives have been saved |
| 2) More than 139,000 | b) the number of beaches the lifeguards watch       |
| 3) More than 4,500   | c) the number of seasonal lifeguards                |
| 4) More than 900     | d) the number of people working on lifeboats        |

**D. Read *RNLI volunteers*. Are they about Robin Howell (R) or Ben Ellington (B)?**

- 1) \_\_\_ He sometimes finds it hard to find free time.
- 2) \_\_\_ He has worked for the organization for eight years.
- 3) \_\_\_ While working, he maintains contact with the lifeguard base.
- 4) \_\_\_ He loves working for the RNLI because he wants to help the community.

**AFTER READING****E. Discuss the following question with a partner.**

Would you like to do any volunteer work? If no, why not? If yes, what sort of work would you like to do?



## UPPER-INTERMEDIATE LEVEL:

### BEFORE READING

#### A. Discuss the question below in class.

What makes a city popular...?

- among its residents
- among visitors

### WHILE READING

#### *The Best City in the World*

Whenever surveys listing the most liveable cities in the world are published, the Canadian city of Vancouver is usually right there at the top of the rankings. There is evidence that people have lived in the Vancouver area for over 10,000 years. With a plentiful supply of food, it was the perfect place for a settlement. Europeans started arriving in the area in the mid-18th century. Today, Vancouver is home to over 2.3 million people and has over 8.5 million visitors a year. What makes the city such a special place?

Vancouver has everything like the skyscrapers, the nature and the sea. There is a kind of friendliness to this city where people seem open to helping each other out and talking to strangers. Vancouver has an abundance of things to do. There are many parks, mountains and different outdoor activities to do. Vancouver's natural beauty is impossible to ignore. From snow-capped mountain peaks to lush forests and beautiful bays and beaches, the scenery is stunning.

Vancouverites are passionate about their home team, the Canucks. The team have yet to win the Stanley Cup, the ultimate prize in ice-hockey, but their dedicated fans will never give up hope. Besides ice-hockey, you can do sailing or kayaking on the rivers and creeks around the city.

120-year-old Stanley Park in the city is one of the biggest and most dramatic urban parks in the world. It boasts a host of attractions, and the park is also famous for its First Nation art. No visit is complete without a photo of the remarkable totem pole display at Brockton Point.

Vancouver is a diverse and cosmopolitan city. Around half of its residents are immigrants to Canada, with 25% considering Chinese as their native tongue. With this diversity comes many benefits, not least the amazing variety of delicious cuisine on offer in the city.

#### *Greenest City 2020*

Vancouver is not complacent about the challenges that come with being one of the world's most desirable destinations, and it has an action plan to become the greenest city in the world in 2020. The Greenest City Plan was put together after consultations online with more than 35,000 residents. The plan called on all the residents to rise to the challenge of transforming the community to create a better life for future generations. It sets out ten achievable goals for the next decade. Notable goals include plans to reduce greenhouse gas emissions by 33%, to help residents make the majority of trips by foot, to plant 150,000 new trees and to ensure that no resident lives more than five minutes away from a green space.

Gwendal Castellan works for the city helping to implement the plan. He explains that the Greenest City Action Plan is a plan that addresses many different goals to make the city more sustainable, so people will see that Vancouver is one that is really trying to be a city that has a lower footprint. But the economy of Vancouver is dependent on tourism, and tourists don't always make the most environmentally friendly of guests. Gwendal believes that the city's plans could have a far-reaching effect on its visitors. He thinks once you have experienced something that is different, that really shows you know better practice or more conscientious stewardship of the environment.

Vancouver will need the support of all its residents and visitors to achieve its Greenest City Plan. It won't be easy, but just one look at the stunning city makes you believe that preserving this place for future generations will be worth all the work.

**B. Read *The Best City in the World*. Underline the correct answers.**

- 1) Vancouver is *usually* / *always* voted the best place to live.
- 2) The settlement of Vancouver dates back *over 10,000 years* / *to the 18<sup>th</sup> century*.
- 3) Vancouver *offers its visitors a wide range of things to do* / *is not very hospitable to newcomers*.
- 4) Vancouver's ice-hockey team *often wins the Stanley Cup* / *is strongly supported by its fans*.
- 5) Stanley Park is *used as an art gallery* / *home to some famous artworks*.

**C. Read *Greenest City 2020*. Complete the sentences with a word or short phrase from the text.**

- 1) The Greenest City Action Plan \_\_\_\_\_ ten main objectives for the next ten years.
- 2) One goal is to make sure that everyone lives close to a \_\_\_\_\_.
- 3) Gwendal hopes that the plan will help the city be more \_\_\_\_\_.
- 4) Vancouver is economically \_\_\_\_\_ tourism.
- 5) The plan aims to \_\_\_\_\_ the city for future generations.

**AFTER READING**

**D. Discuss the following question with a partner.**

What do you think of Vancouver's plan? Are their goals achievable?

## APPENDIX B

### WATCHING-ONLY (WO) TASKS

#### LOWER-INTERMEDIATE LEVEL:

##### BEFORE WATCHING

##### A. Discuss the question below in class.

The sea can be a dangerous place. What risks are there for...?

- surfers and swimmers
- fishermen and sailors

##### WHILE WATCHING

##### B. Watch *Free time by the sea*. Make notes to answer the questions.

- 1) What relaxing activities can you do in Cornwall?
- 2) Name three active sports you can do there.
- 3) What are some of the dangers of the Cornish coast?

##### C. Watch *The RNLI*. Match the numbers to what they refer to.

- |                      |                                                     |
|----------------------|-----------------------------------------------------|
| 1) 160               | a) the number of people whose lives have been saved |
| 2) More than 139,000 | b) the number of beaches the lifeguards watch       |
| 3) More than 4,500   | c) the number of seasonal lifeguards                |
| 4) More than 900     | d) the number of people working on lifeboats        |

##### D. Watch *RNLI volunteers*. Are they about Robin Howell (R) or Ben Ellington (B)?

- 1) \_\_\_ He sometimes finds it hard to find free time.
- 2) \_\_\_ He has worked for the organization for eight years.
- 3) \_\_\_ While working, he maintains contact with the lifeguard base.
- 4) \_\_\_ He loves working for the RNLI because he wants to help the community.

##### AFTER WATCHING

##### E. Discuss the following question with a partner.

Would you like to do any volunteer work? If no, why not? If yes, what sort of work would you like to do?

**UPPER-INTERMEDIATE LEVEL:****BEFORE WATCHING****A. Discuss the question below in class.**

What makes a city popular...?

- among its residents
- among visitors

**WHILE WATCHING****B. Watch *The Best City in the World*. Underline the correct answers.**

- 1) Vancouver is ***usually*** / ***always*** voted the best place to live.
- 2) The settlement of Vancouver dates back ***over 10,000 years*** / ***to the 18<sup>th</sup> century***.
- 3) Vancouver ***offers its visitors a wide range of things to do*** / ***is not very hospitable to newcomers***.
- 4) Vancouver's ice-hockey team ***often wins the Stanley Cup*** / ***is strongly supported by its fans***.
- 5) Stanley Park is ***used as an art gallery*** / ***home to some famous artworks***.

**C. Watch *Greenest City 2020*. Complete the sentences with a word or short phrase from the video.**

- 1) The Greenest City Action Plan \_\_\_\_\_ ten main objectives for the next ten years.
- 2) One goal is to make sure that everyone lives close to a \_\_\_\_\_.
- 3) Gwendal hopes that the plan will help the city be more \_\_\_\_\_.
- 4) Vancouver is economically \_\_\_\_\_ tourism.
- 5) The plan aims to \_\_\_\_\_ the city for future generations.

**AFTER WATCHING****D. Discuss the following question with a partner.**

What do you think of Vancouver's plan? Are their goals achievable?

## APPENDIX C

### FILL-IN-BY-READING (FR) TASKS

#### LOWER-INTERMEDIATE LEVEL:

##### BEFORE READING

##### A. Discuss the question below in class.

The sea can be a dangerous place. What risks are there for...?

- surfers and swimmers
- fishermen and sailors

##### WHILE READING

##### B. Read the text below. Complete the blanks with the words given while you are reading.

**pager (n.):** an electronic device that beeps when people contact you (çağrı cihazı)

**treacherous (adj.):** extremely dangerous (tehlikeli)

**dedicated (adj.):** giving a lot of time/energy to something important (kendini işine adanmış)

**stunning (adj.):** very beautiful (büyüleyici)

**patrol (v.):** to go around an area to see if there is any danger (devriye gezmek)

**current(s) (n.):** a movement of water in a particular direction (akıntı)

**pupil(s) (n.):** a student (öğrenci)

**launch (v.):** to set (a boat or ship) in the water (denize indirmek)

#### The Working Week

##### *Free time by the sea*

Cornwall is in the south-west of England. It is famous for its \_\_\_\_\_ scenery, pretty villages and beautiful beaches. And there's plenty to do with your free time here. For some people, the weekend is about taking the time to relax. At the beach, people take their dogs for walks, catch up with friends or spend time with their families. And for others, it's all about getting active. You can be a successful accountant during the week and a passionate surfer at the weekend. You can go for a swim or perhaps explore the coastline by kayak. But while the sea and coastline are beautiful, they can also be \_\_\_\_\_, with strong \_\_\_\_\_ and unpredictable weather conditions.

##### **The RNLI**

Every year, thousands of people get into difficulty on the coast, and it's not just the holiday-makers. In the UK, when something goes wrong at sea or on the beach, there's an amazing group of \_\_\_\_\_ people who give up their free time to help: the members of the Royal National Lifeboat Institution, the RNLI. It is a charity that saves lives at sea. The RNLI was formed in 1824 to help anyone who got into trouble in the seas around the coastline of the UK. It has both lifeboat volunteers who respond to emergency calls, and lifeguards who \_\_\_\_\_ more than 160 beaches. The charity uses money donated by the public to pay for its work.

Since it started, the RNLI's boats and lifeguards have saved more than 139,000 lives. Today, there are over four and a half thousand lifeboat crew and over 900 seasonal lifeguards, working for the organization. Most of these people are volunteers who put their own lives in danger to prevent accidents and rescue people in need. For RNLI lifeguards, 95 per cent of the job is to prevent accidents before they happen at the beach. They watch the sea carefully to check that no one is in trouble. They only stay on duty for short periods of time before they have to take a break. Lifeboat crew members attend emergencies at sea. Their \_\_\_\_\_ will go off, and they will rush down to the lifeboat station and \_\_\_\_\_ the lifeboat.

### ***RNLI volunteers***

What kind of people volunteer for the RNLI? Robin Howell is a senior lifeguard on Perranporth beach in Cornwall. It's 10 a.m. on a Sunday in June, and Robin is starting work. He's setting up all his equipment and getting ready for a day keeping people safe on the beach. He has a radio to stay in contact with the lifeguard base. Robin has another job. He is a geography teacher at a secondary school in Cornwall. He works all year round as a teacher. As a teacher, he doesn't work at the weekends, but from March all the way through to September he works as a lifeguard at the weekends. His lifeguard hours are between 10 and 6 p.m. in the evening.

Having two jobs doesn't leave much time for relaxing, and even when he's not working, Robin likes to stay busy. He says, "I think the main problem with having two jobs is just literally trying to fit it in the space of a week, but I'd rather be busy than sitting at home watching TV. There's not a lot of free time, but I try to do as much sport as possible. I swim in the mornings at 6.30. I have two jobs because I love both of them. I love the lifeguarding because it's outside, it's very physical and I love the teaching because I just like the interaction with the \_\_\_\_\_ and the fact that I can try and make a difference to their education."

25 year-old Ben Eglington is part of the RNLI lifeboat crew in the town of Newquay in Cornwall. His crew is made up of many volunteers from the local community. There are fishermen, builders, landscape gardeners, hotel/pub owners and fishmongers. Ben's been part of the crew for eight years. He's a busy man, who works full-time as an engineer in the Royal Navy during the week. During the week, he works away in Portsmouth, which is 210 miles away. Every weekend, he makes the epic journey back to Newquay to help out with the RNLI, but Ben doesn't mind giving up his free time to help. He says, "I volunteer my time with the lifeboat because it's good to give something back to the community. It's also a great achievement when you're out there and you actually save somebody's life."

It's thanks to Robin, Ben and all the amazing RNLI volunteers that the beautiful beaches, coastline and seas around the UK are safer for everyone to enjoy.

### **C. Read *Free time by the sea*. Make notes to answer the questions.**

- 1) What relaxing activities can you do in Cornwall?
- 2) Name three active sports you can do there.
- 3) What are some of the dangers of the Cornish coast?

**D. Read *The RNLI*. Match the numbers to what they refer to.**

- |                      |                                                     |
|----------------------|-----------------------------------------------------|
| 1) 160               | a) the number of people whose lives have been saved |
| 2) More than 139,000 | b) the number of beaches the lifeguards watch       |
| 3) More than 4,500   | c) the number of seasonal lifeguards                |
| 4) More than 900     | d) the number of people working on lifeboats        |

**E. Read *RNLI volunteers*. Are they about Robin Howell (R) or Ben Ellington (B)?**

- 1) \_\_\_ He sometimes finds it hard to find free time.
- 2) \_\_\_ He has worked for the organization for eight years.
- 3) \_\_\_ While working, he maintains contact with the lifeguard base.
- 4) \_\_\_ He loves working for the RNLI because he wants to help the community.

**AFTER READING****F. Discuss the following question with a partner.**

Would you like to do any volunteer work? If no, why not? If yes, what sort of work would you like to do?

## UPPER-INTERMEDIATE LEVEL:

### BEFORE READING

#### A. Discuss the question below in class.

What makes a city popular...?

- among its residents
- among visitors

### WHILE READING

#### B. Read the text below. Complete the blanks with the words given while you are reading.

**display (n.):** a public exhibition (sergi)

**conscientious (adj.):** concerned with doing something right (insaflı)

**ultimate (adj.):** the highest and most important (en üst ve önemli)

**complacent (adj.):** satisfied with what you have (elindekiyle yetinen)

**implement (v.):** to put into practice (uygulamaya koymak)

**creek(s) (n.):** a narrow waterway (ırmak kolu)

**abundance (n.):** a great amount (büyük miktar)

**boast (v.):** to be proud of having something (sahip olmakla övünmek)

#### *The Best City in the World*

Whenever surveys listing the most liveable cities in the world are published, the Canadian city of Vancouver is usually right there at the top of the rankings. There is evidence that people have lived in the Vancouver area for over 10,000 years. With a plentiful supply of food, it was the perfect place for a settlement. Europeans started arriving in the area in the mid-18th century. Today, Vancouver is home to over 2.3 million people and has over 8.5 million visitors a year. What makes the city such a special place?

Vancouver has everything like the skyscrapers, the nature and the sea. There is a kind of friendliness to this city where people seem open to helping each other out and talking to strangers. Vancouver has a(n) \_\_\_\_\_ of things to do. There are many parks, mountains and different outdoor activities to do. Vancouver's natural beauty is impossible to ignore. From snow-capped mountain peaks to lush forests and beautiful bays and beaches, the scenery is stunning.

Vancouverites are passionate about their home team, the Canucks. The team have yet to win the Stanley Cup, the \_\_\_\_\_ prize in ice-hockey, but their dedicated fans will never give up hope. Besides ice-hockey, you can do sailing or kayaking on the rivers and \_\_\_\_\_ around the city.

120-year-old Stanley Park in the city is one of the biggest and most dramatic urban parks in the world. It \_\_\_\_\_ a host of attractions, and the park is also famous for its First Nation art. No visit is complete without a photo of the remarkable totem pole \_\_\_\_\_ at Brockton Point.

Vancouver is a diverse and cosmopolitan city. Around half of its residents are immigrants to Canada, with 25% considering Chinese as their native tongue. With this diversity comes many benefits, not least the amazing variety of delicious cuisine on offer in the city.

### **Greenest City 2020**

Vancouver is not \_\_\_\_\_ about the challenges that come with being one of the world's most desirable destinations, and it has an action plan to become the greenest city in the world in 2020. The Greenest City Plan was put together after consultations online with more than 35,000 residents. The plan called on all the residents to rise to the challenge of transforming the community to create a better life for future generations. It sets out ten achievable goals for the next decade. Notable goals include plans to reduce greenhouse gas emissions by 33%, to help residents make the majority of trips by foot, to plant 150,000 new trees and to ensure that no resident lives more than five minutes away from a green space.

Gwendal Castellán works for the city helping to \_\_\_\_\_ the plan. He explains that the Greenest City Action Plan is a plan that addresses many different goals to make the city more sustainable, so people will see that Vancouver is one that is really trying to be a city that has a lower footprint. But the economy of Vancouver is dependent on tourism, and tourists don't always make the most environmentally friendly of guests. Gwendal believes that the city's plans could have a far-reaching effect on its visitors. He thinks once you have experienced something that is different, that really shows you know better practice or more \_\_\_\_\_ stewardship of the environment.

Vancouver will need the support of all its residents and visitors to achieve its Greenest City Plan. It won't be easy, but just one look at the stunning city makes you believe that preserving this place for future generations will be worth all the work.

### **C. Read *The Best City in the World*. Underline the correct answers.**

- 1) Vancouver is *usually* / *always* voted the best place to live.
- 2) The settlement of Vancouver dates back *over 10,000 years* / *to the 18<sup>th</sup> century*.
- 3) Vancouver *offers its visitors a wide range of things to do* / *is not very hospitable to newcomers*.
- 4) Vancouver's ice-hockey team *often wins the Stanley Cup* / *is strongly supported by its fans*.
- 5) Stanley Park is *used as an art gallery* / *home to some famous artworks*.

### **D. Read *Greenest City 2020*. Complete the sentences with a word or short phrase from the text.**

- 1) The Greenest City Action Plan \_\_\_\_\_ ten main objectives for the next ten years.
- 2) One goal is to make sure that everyone lives close to a \_\_\_\_\_.
- 3) Gwendal hopes that the plan will help the city be more \_\_\_\_\_.
- 4) Vancouver is economically \_\_\_\_\_ tourism.
- 5) The plan aims to \_\_\_\_\_ the city for future generations.

### **AFTER READING**

#### **E. Discuss the following question with a partner.**

What do you think of Vancouver's plan? Are their goals achievable?

## APPENDIX D

### FILL-IN-BY-WATCHING (FW) TASKS

#### LOWER-INTERMEDIATE LEVEL:

##### BEFORE WATCHING

##### A. Discuss the question below in class.

The sea can be a dangerous place. What risks are there for...?

- surfers and swimmers
- fishermen and sailors

##### WHILE WATCHING

##### B. The sentences below are extracted from the video *The Working Week*. Complete the blanks with the words given while you are watching.

**pager (n.):** an electronic device that beeps when people contact you (çağrı cihazı)

**treacherous (adj.):** extremely dangerous (tehlikeli)

**dedicated (adj.):** giving a lot of time/energy to something important (kendini işine adanmış)

**stunning (adj.):** very beautiful (büyüleyici)

**patrol (v.):** to go around an area to see if there is any danger (devriye gezmek)

**current(s) (n.):** a movement of water in a particular direction (akıntı)

**pupil(s) (n.):** a student (öğrenci)

**launch (v.):** to set (a boat or ship) in the water (denize indirmek)

- 1) Cornwall is famous for its \_\_\_\_\_ scenery, pretty villages and beautiful beaches.
- 2) But while the sea and coastline are beautiful, they can also be \_\_\_\_\_, with strong \_\_\_\_\_ and unpredictable weather conditions.
- 3) In the UK, when something goes wrong at sea or on the beach..., there's an amazing group of \_\_\_\_\_ people who give up their free time to help.
- 4) The RNLI has both lifeboat volunteers who respond to emergency calls 24 hours a day, and lifeguards who \_\_\_\_\_ more than 160 beaches.
- 5) The crew members' \_\_\_\_\_ will go off, and they will rush down to the lifeboat station and \_\_\_\_\_ the lifeboat.
- 6) I love the teaching because I just like the interaction with the \_\_\_\_\_.

##### C. Watch *Free time by the sea*. Make notes to answer the questions.

- 1) What relaxing activities can you do in Cornwall?
- 2) Name three active sports you can do there.
- 3) What are some of the dangers of the Cornish coast?

**D. Watch *The RNLI*. Match the numbers to what they refer to.**

- |                      |                                                     |
|----------------------|-----------------------------------------------------|
| 1) 160               | a) the number of people whose lives have been saved |
| 2) More than 139,000 | b) the number of beaches the lifeguards watch       |
| 3) More than 4,500   | c) the number of seasonal lifeguards                |
| 4) More than 900     | d) the number of people working on lifeboats        |

**E. Watch *RNLI volunteers*. Are they about Robin Howell (R) or Ben Ellington (B)?**

- 1) \_\_\_ He sometimes finds it hard to find free time.
- 2) \_\_\_ He has worked for the organization for eight years.
- 3) \_\_\_ While working, he maintains contact with the lifeguard base.
- 4) \_\_\_ He loves working for the RNLI because he wants to help the community.

**AFTER WATCHING****F. Discuss the following question with a partner.**

Would you like to do any volunteer work? If no, why not? If yes, what sort of work would you like to do?

## UPPER-INTERMEDIATE LEVEL:

### BEFORE WATCHING

#### A. Discuss the question below in class.

What makes a city popular...?

- among its residents
- among visitors

### WHILE WATCHING

#### B. The sentences below are extracted from the video *Vancouver*. Complete the blanks with the words given while you are watching.

**display (n.):** a public exhibition (sergi)

**conscientious (adj.):** concerned with doing something right (insafli)

**ultimate (adj.):** the highest and most important (en üst ve önemli)

**complacent (adj.):** satisfied with what you have (elindekiyle yetinen)

**implement (v.):** to put into practice (uygulamaya koymak)

**creek(s) (n.):** a narrow waterway (ırmak kolu)

**abundance (n.):** a great amount (büyük miktar)

**boast (v.):** to be proud of having something (sahip olmakla övünmek)

- 1) Vancouver just has a(n) \_\_\_\_\_ of things to do, kind of outside of your work life...
- 2) Vancouverites are passionate about their home team, the Canucks. The team have yet to win the Stanley Cup, the \_\_\_\_\_ prize in ice-hockey.
- 3) You can do sailing or kayaking on the rivers and \_\_\_\_\_ around the city.
- 4) The 120-year-old green space \_\_\_\_\_ a host of attractions, and the park is also famous for its First Nation art, and no visit is complete without a photo of the remarkable totem pole \_\_\_\_\_ at Brockton Point.
- 5) Vancouver is not \_\_\_\_\_ about the challenges that come with being one of the world's most desirable destinations.
- 6) Gwendal Castellan works for the city of Vancouver helping to \_\_\_\_\_ the plan.
- 7) One area where I really see the opportunity for a global change is the idea that once you've experienced something that's different, that really shows, you know, better practice or \_\_\_\_\_ stewardship of the environment.

#### C. Watch *The Best City in the World*. Underline the correct answers.

- 1) Vancouver is usually / always voted the best place to live.
- 2) The settlement of Vancouver dates back over 10,000 years / to the 18<sup>th</sup> century.
- 3) Vancouver offers its visitors a wide range of things to do / is not very hospitable to newcomers.
- 4) Vancouver's ice-hockey team often wins the Stanley Cup / is strongly supported by its fans.
- 5) Stanley Park is used as an art gallery / home to some famous artworks.

**D. Watch *Greenest City 2020*. Complete the sentences with a word or short phrase from the video.**

- 1) The Greenest City Action Plan \_\_\_\_\_ ten main objectives for the next ten years.
- 2) One goal is to make sure that everyone lives close to a \_\_\_\_\_.
- 3) Gwendal hopes that the plan will help the city be more \_\_\_\_\_.
- 4) Vancouver is economically \_\_\_\_\_ tourism.
- 5) The plan aims to \_\_\_\_\_ the city for future generations.

**AFTER WATCHING**

**E. Discuss the following question with a partner.**

What do you think of Vancouver's plan? Are their goals achievable?

## APPENDIX E

### SENTENCE-WRITING-BY-READING (SR) TASKS

#### LOWER-INTERMEDIATE LEVEL:

##### BEFORE READING

##### A. Discuss the question in class.

The sea can be a dangerous place. What risks are there for...?

- surfers and swimmers
- fishermen and sailors

##### WHILE READING

##### B. Read the text below and look at the definitions of the bold words while you are reading.

**stunning (adj.):** very beautiful (büyüleyici)

**treacherous (adj.):** extremely dangerous (tehlikeli)

**current(s) (n.):** a movement of water in a particular direction (akıntı)

**dedicated (adj.):** giving a lot of time/energy to something important (kendini işine adanmış)

**patrol (v.):** to go around an area to see if there is any danger (devriye gezmek)

**pager (n.):** an electronic device that beeps when people contact you (çağrı cihazı)

**launch (v.):** to set (a boat or ship) in the water (denize indirmek)

**pupil(s) (n.):** a student (öğrenci)

#### The Working Week

##### *Free time by the sea*

Cornwall is in the south-west of England. It is famous for its **stunning** scenery, pretty villages and beautiful beaches. And there's plenty to do with your free time here. For some people, the weekend is about taking the time to relax. At the beach, people take their dogs for walks, catch up with friends or spend time with their families. And for others, it's all about getting active. You can be a successful accountant during the week and a passionate surfer at the weekend. You can go for a swim or perhaps explore the coastline by kayak. But while the sea and coastline are beautiful, they can also be **treacherous**, with strong **currents** and unpredictable weather conditions.

##### *The RNLI*

Every year, thousands of people get into difficulty on the coast, and it's not just the holiday-makers. In the UK, when something goes wrong at sea or on the beach, there's an amazing group of **dedicated** people who give up their free time to help: the members of the Royal National Lifeboat Institution, the RNLI. It is a charity that saves lives at sea. The RNLI was formed in 1824 to help anyone who got into trouble in the seas around the coastline of the UK. It has both lifeboat volunteers who respond to emergency calls, and lifeguards who **patrol** more than 160 beaches. The charity uses money donated by the public to pay for its work.

Since it started, the RNLI's boats and lifeguards have saved more than 139,000 lives.

Today, there are over four and a half thousand lifeboat crew and over 900 seasonal lifeguards, working for the organization. Most of these people are volunteers who put their own lives in danger to prevent accidents and rescue people in need. For RNLI lifeguards, 95 per cent of the job is to prevent accidents before they happen at the beach. They watch the sea carefully to check that no one is in trouble. They only stay on duty for short periods of time before they have to take a break. Lifeboat crew members attend emergencies at sea. Their **pager** will go off, and they will rush down to the lifeboat station and **launch** the lifeboat.

### ***RNLI volunteers***

What kind of people volunteer for the RNLI? Robin Howell is a senior lifeguard on Perranporth beach in Cornwall. It's 10 a.m. on a Sunday in June, and Robin is starting work. He's setting up all his equipment and getting ready for a day keeping people safe on the beach. He has a radio to stay in contact with the lifeguard base. Robin has another job. He is a geography teacher at a secondary school in Cornwall. He works all year round as a teacher. As a teacher, he doesn't work at the weekends, but from March all the way through to September he works as a lifeguard at the weekends. His lifeguard hours are between 10 and 6 p.m. in the evening.

Having two jobs doesn't leave much time for relaxing, and even when he's not working, Robin likes to stay busy. He says, "I think the main problem with having two jobs is just literally trying to fit it in the space of a week, but I'd rather be busy than sitting at home watching TV. There's not a lot of free time, but I try to do as much sport as possible. I swim in the mornings at 6.30. I have two jobs because I love both of them. I love the lifeguarding because it's outside, it's very physical and I love the teaching because I just like the interaction with the **pupils** and the fact that I can try and make a difference to their education."

25 year-old Ben Eglington is part of the RNLI lifeboat crew in the town of Newquay in Cornwall. His crew is made up of many volunteers from the local community. There are fishermen, builders, landscape gardeners, hotel/pub owners and fishmongers. Ben's been part of the crew for eight years. He's a busy man, who works full-time as an engineer in the Royal Navy during the week. During the week, he works away in Portsmouth, which is 210 miles away. Every weekend, he makes the epic journey back to Newquay to help out with the RNLI, but Ben doesn't mind giving up his free time to help. He says, "I volunteer my time with the lifeboat because it's good to give something back to the community. It's also a great achievement when you're out there and you actually save somebody's life."

It's thanks to Robin, Ben and all the amazing RNLI volunteers that the beautiful beaches, coastline and seas around the UK are safer for everyone to enjoy.

### **C. Read *Free time by the sea*. Make notes to answer the questions.**

- 1) What relaxing activities can you do in Cornwall?
- 2) Name three active sports you can do there.
- 3) What are some of the dangers of the Cornish coast?

**D. Read *RNLI volunteers*. Are they about Robin Howell (R) or Ben Ellington (B)?**

- 1) \_\_\_ He sometimes finds it hard to find free time.
- 2) \_\_\_ He has worked for the organization for eight years.
- 3) \_\_\_ While working, he maintains contact with the lifeguard base.
- 4) \_\_\_ He loves working for the RNLI because he wants to help the community.

**AFTER READING**

**E. Write an example sentence for each word below. Your sentences should help understand the meaning of each word.**

1. **stunning (adj.):** very beautiful (büyüleyici)

2. **treacherous (adj.):** extremely dangerous (tehlikeli)

3. **current(s) (n.):** a movement of water in a particular direction (akıntı)

4. **dedicated (adj.):** giving a lot of time/energy to something important (kendini işine adanmış)

5. **patrol (v.):** to go around an area to see if there is any danger (devriye gezmek)

6. **pager (n.):** an electronic device that beeps when people contact you (çağrı cihazı)

7. **launch (v.):** to set (a boat or ship) in the water (denize indirmek)

8. **pupil(s) (n.):** a student (öğrenci)

**F. Discuss the following question with a partner.**

Would you like to do any volunteer work? If no, why not? If yes, what sort of work would you like to do?

## UPPER-INTERMEDIATE LEVEL:

### BEFORE READING

#### A. Discuss the question below in class.

What makes a city popular...?

- among its residents
- among visitors

### WHILE READING

#### B. Read the text below and look at the definitions of the bold words while you are reading.

**abundance (n.):** a great amount (büyük miktar)

**ultimate (adj.):** the highest and most important (en üst ve önemli)

**creek(s) (n.):** a narrow waterway (ırmak kolu)

**boast (v.):** to be proud of having something (sahip olmakla övünmek)

**display (n.):** a public exhibition (sergi)

**complacent (adj.):** satisfied with what you have (elindekiyle yetinen)

**implement (v.):** to put into practice (uygulamaya koymak)

**conscientious (adj.):** concerned with doing something right (insaflı)

#### The Best City in the World

Whenever surveys listing the most liveable cities in the world are published, the Canadian city of Vancouver is usually right there at the top of the rankings. There is evidence that people have lived in the Vancouver area for over 10,000 years. With a plentiful supply of food, it was the perfect place for a settlement. Europeans started arriving in the area in the mid-18th century. Today, Vancouver is home to over 2.3 million people and has over 8.5 million visitors a year. What makes the city such a special place?

Vancouver has everything like the skyscrapers, the nature and the sea. There is a kind of friendliness to this city where people seem open to helping each other out and talking to strangers. Vancouver has an **abundance** of things to do. There are many parks, mountains and different outdoor activities to do. Vancouver's natural beauty is impossible to ignore. From snow-capped mountain peaks to lush forests and beautiful bays and beaches, the scenery is stunning.

Vancouverites are passionate about their home team, the Canucks. The team have yet to win the Stanley Cup, the **ultimate** prize in ice-hockey, but their dedicated fans will never give up hope. Besides ice-hockey, you can do sailing or kayaking on the rivers and **creeks** around the city.

120-year-old Stanley Park in the city is one of the biggest and most dramatic urban parks in the world. It **boasts** a host of attractions, and the park is also famous for its First Nation art. No visit is complete without a photo of the remarkable totem pole **display** at Brockton Point.

Vancouver is a diverse and cosmopolitan city. Around half of its residents are immigrants to Canada, with 25% considering Chinese as their native tongue. With this diversity comes many benefits, not least the amazing variety of delicious cuisine on offer in the city.

### Greenest City 2020

Vancouver is not **complacent** about the challenges that come with being one of the world's most desirable destinations, and it has an action plan to become the greenest city in the world in 2020. The Greenest City Plan was put together after consultations online with more than 35,000 residents. The plan called on all the residents to rise to the challenge of transforming the community to create a better life for future generations. It sets out ten achievable goals for the next decade. Notable goals include plans to reduce greenhouse gas emissions by 33%, to help residents make the majority of trips by foot, to plant 150,000 new trees and to ensure that no resident lives more than five minutes away from a green space.

Gwendal Castellan works for the city helping to **implement** the plan. He explains that the Greenest City Action Plan is a plan that addresses many different goals to make the city more sustainable, so people will see that Vancouver is one that is really trying to be a city that has a lower footprint. But the economy of Vancouver is dependent on tourism, and tourists don't always make the most environmentally friendly of guests. Gwendal believes that the city's plans could have a far-reaching effect on its visitors. He thinks once you have experienced something that is different, that really shows you know better practice or more **conscientious** stewardship of the environment.

Vancouver will need the support of all its residents and visitors to achieve its Greenest City Plan. It won't be easy, but just one look at the stunning city makes you believe that preserving this place for future generations will be worth all the work.

### C. Read *The Best City in the World*. Underline the correct answers.

- 1) Vancouver is *usually* / *always* voted the best place to live.
- 2) The settlement of Vancouver dates back *to the 18<sup>th</sup> century* / *over 10,000 years*.
- 3) Vancouver *offers its visitors a wide range of things to do* / *is not very hospitable to newcomers*.

### D. Read *Greenest City 2020*. Complete the sentences with a short phrase from the text.

- 1) The Greenest City Action Plan \_\_\_\_\_ ten main objectives for the next ten years.
- 2) One goal is to make sure that everyone lives close to a \_\_\_\_\_.
- 3) Vancouver is economically \_\_\_\_\_ tourism.

### AFTER READING

### E. Write an example sentence for each word below. Your sentences should help understand the meaning of each word.

1. **abundance (n.):** a great amount (büyük miktar)

\_\_\_\_\_

2. **ultimate (adj.):** the highest and most important (en üst ve önemli)

\_\_\_\_\_

3. **creek(s) (n.):** a narrow waterway (ırmak kolu)

\_\_\_\_\_

4. **boast (v.):** to be proud of having something (sahip olmakla övünmek)

\_\_\_\_\_

5. **display (n.):** a public exhibition (sergi)

---

6. **complacent (adj.):** satisfied with what you have (elindekiyle yetinen)

---

7. **implement (v.):** to put into practice (uygulamaya koymak)

---

8. **conscientious (adj.):** attentive and careful (özenli ve dikkatli)

---

**F. Discuss the following question with a partner.**

What do you think of Vancouver's plan? Are their goals achievable?

## APPENDIX F

### SENTENCE-WRITING-BY-WATCHING (SW) TASKS

#### LOWER-INTERMEDIATE LEVEL:

##### BEFORE WATCHING

###### A. Discuss the question in class.

The sea can be a dangerous place. What risks are there for...?

- surfers and swimmers
- fishermen and sailors

###### B. The words below are extracted from the video *The Working Week*. Look at their definitions before watching.

**stunning (adj.):** very beautiful (büyüleyici)

**treacherous (adj.):** extremely dangerous (tehlikeli)

**current(s) (n.):** a movement of water in a particular direction (akıntı)

**dedicated (adj.):** giving a lot of time/energy to something important (kendini işine adanmış)

**patrol (v.):** to go around an area to see if there is any danger (devriye gezmek)

**pager (n.):** an electronic device that beeps when people contact you (çağrı cihazı)

**launch (v.):** to set (a boat or ship) in the water (denize indirmek)

**pupil(s) (n.):** a student (öğrenci)

##### WHILE WATCHING

###### C. Watch *Free time by the sea*. Make notes to answer the questions.

- 1) What relaxing activities can you do in Cornwall?
- 2) Name three active sports you can do there.
- 3) What are some of the dangers of the Cornish coast?

###### D. Watch *RNLI volunteers*. Are they about Robin Howell (R) or Ben Ellington (B)?

- 1) \_\_\_ He sometimes finds it hard to find free time.
- 2) \_\_\_ He has worked for the organization for eight years.
- 3) \_\_\_ While working, he maintains contact with the lifeguard base.
- 4) \_\_\_ He loves working for the RNLI because he wants to help the community.

##### AFTER WATCHING

###### E. Write an example sentence for each word below. Your sentences should help understand the meaning of each word.

1. **stunning (adj.):** very beautiful (büyüleyici)

2. **treacherous (adj.):** extremely dangerous (tehlikeli)

3. **current(s) (n.):** a movement of water in a particular direction (akıntı)

**4. dedicated (adj.):** giving a lot of time/energy to something important (kendini işine adanmış)

---

**5. patrol (v.):** to go around an area to see if there is any danger (devriye gezmek)

---

**6. pager (n.):** an electronic device that beeps when people contact you (çağrı cihazı)

---

**7. launch (v.):** to set (a boat or ship) in the water (denize indirmek)

---

**8. pupil(s) (n.):** a student (öğrenci)

---

**F. Discuss the following question with a partner.**

Would you like to do any volunteer work? If no, why not? If yes, what sort of work would you like to do?

## UPPER-INTERMEDIATE LEVEL:

### BEFORE WATCHING

#### A. Discuss the question below in class.

What makes a city popular...?

- among its residents
- among tourists

#### B. The words below are extracted from the video *Vancouver*. Look at their definitions before watching.

**abundance (n.):** a great amount (büyük miktar)

**ultimate (adj.):** the highest and most important (en üst ve önemli)

**creek(s) (n.):** a narrow waterway (ırmak kolu)

**boast (v.):** to be proud of having something (sahip olmakla övünmek)

**display (n.):** a public exhibition (sergi)

**complacent (adj.):** satisfied with what you have (elindekiyle yetinen)

**implement (v.):** to put into practice (uygulamaya koymak)

**conscientious (adj.):** concerned with doing something right (insafli)

### WHILE WATCHING

#### C. Watch *The Best City in the World*. Underline the correct answers.

- 1) Vancouver is **usually** / **always** voted the best place to live.
- 2) The settlement of Vancouver dates back **to the 18<sup>th</sup> century** / **over 10,000 years**.
- 3) Vancouver **offers its visitors a wide range of things to do** / **is not very hospitable to newcomers**.

#### D. Watch *Greenest City 2020*. Complete the sentences with a short phrase from the video.

- 1) The Greenest City Action Plan \_\_\_\_\_ ten main objectives for the next ten years.
- 2) One goal is to make sure that everyone lives close to a \_\_\_\_\_.
- 3) Vancouver is economically \_\_\_\_\_ tourism.

### AFTER WATCHING

#### E. Write an example sentence for each word below. Your sentences should help understand the meaning of each word.

1. **abundance (n.):** a great amount (büyük miktar)

2. **ultimate (adj.):** the highest and most important (en üst ve önemli)

3. **creek(s) (n.):** a narrow waterway (ırmak kolu)

4. **boast (v.):** to be proud of having something (sahip olmakla övünmek)

5. **display (n.):** a public exhibition (sergi)

---

6. **complacent (adj.):** satisfied with what you have (elindekiyle yetinen)

---

7. **implement (v.):** to put into practice (uygulamaya koymak)

---

8. **conscientious (adj.):** attentive and careful (özenli ve dikkatli)

---

**F. Discuss the following question with a partner.**

What do you think of Vancouver's plan? Are their goals achievable?

## **APPENDIX G**

### **VOCABULARY PRE-TESTS**

#### **LOWER-INTERMEDIATE LEVEL:**

**Write a Turkish definition or an English explanation for the meanings of the words below.**

- 1) patrol (v.):**
- 2) increase (v.):**
- 3) duty (n.):**
- 4) attitude (n.):**
- 5) pager (n.):**
- 6) donate (v.):**
- 7) hectic (adj.):**
- 8) achievement (n.):**
- 9) remote (adj.):**
- 10) bargain (v.):**
- 11) stunning (adj.):**
- 12) territory (n.):**
- 13) capture (v.):**
- 14) treacherous (adj.):**
- 15) launch (v.):**
- 16) dedicated (adj.):**
- 17) current (n.):**
- 18) epic (adj.):**
- 19) faith (n.):**
- 20) pupil (n.):**

**UPPER-INTERMEDIATE LEVEL:**

**Write a Turkish definition or an English explanation for the meanings of the words below.**

- 1)** sustainable (adj.):
- 2)** ranking (n.):
- 3)** remarkable (adj.):
- 4)** cuisine (n.):
- 5)** abundance (n.):
- 6)** lush (adj.):
- 7)** display (n.):
- 8)** stewardship (n.):
- 9)** ultimate (adj.):
- 10)** boast (v.):
- 11)** notable (adj.):
- 12)** creek (n.):
- 13)** ensure (v.):
- 14)** complacent (adj.):
- 15)** implement (v.):
- 16)** conscientious (adj.):
- 17)** preserve (v.):
- 18)** supply (n.):
- 19)** emission (n.):
- 20)** consultation (n.):

## APPENDIX H

### PRODUCTIVE VOCABULARY POST-TESTS

#### LOWER-INTERMEDIATE LEVEL:

Write the words that you have learned next to their definitions below.

1. \_\_\_\_\_ (adj.): giving a lot of time/energy to something important (kendini işine adanmış)
2. \_\_\_\_\_ (n.): an electronic device that beeps when people contact you (çağrı cihazı)
3. \_\_\_\_\_ (v.): to set (a boat or ship) in the water (denize indirmek)
4. \_\_\_\_\_ (n.): a movement of water in a particular direction (akıntı)
5. \_\_\_\_\_ (adj.): extremely dangerous (tehlikeli)
6. \_\_\_\_\_ (n.): a student (öğrenci)
7. \_\_\_\_\_ (adj.): very beautiful (büyüleyici)
8. \_\_\_\_\_ (v.): to go around an area to see if there is any danger (devriye gezmek)

#### UPPER-INTERMEDIATE LEVEL:

Write the words that you have learned next to their definitions below.

1. \_\_\_\_\_ (adj.): concerned with doing something right (insafli)
2. \_\_\_\_\_ (n.): a public exhibition (sergi)
3. \_\_\_\_\_ (v.): to put into practice (uygulamaya koymak)
4. \_\_\_\_\_ (n.): a narrow waterway (ırmak kolu)
5. \_\_\_\_\_ (adj.): the highest and most important (en üst ve önemli)
6. \_\_\_\_\_ (n.): a great amount (büyük miktar)
7. \_\_\_\_\_ (adj.): satisfied with what you have (elindekiyle yetinen)
8. \_\_\_\_\_ (v.): to be proud of having something (sahip olmakla övünmek)

## **APPENDIX I**

### **RECEPTIVE VOCABULARY POST-TESTS**

#### **LOWER-INTERMEDIATE LEVEL:**

**Write a Turkish definition or an English explanation for the words below.**

1. pager (n.):
2. treacherous (adj.):
3. pupil (n.):
4. dedicated (adj.):
5. launch (v.):
6. patrol (v.):
7. current (n.):
8. stunning (adj.):

#### **UPPER-INTERMEDIATE LEVEL:**

**Write a Turkish definition or an English explanation for the words below.**

1. creek(s) (n.):
2. complacent (adj.):
3. abundance (n.):
4. conscientious (adj.):
5. boast (v.):
6. implement (v.):
7. display (n.):
8. ultimate (adj.):

## APPENDIX J

### LEXICAL PRODUCTION SCORING PROTOCOL

**Table J. Lexical Production Scoring Protocol**

| .00 points                                                                                                                                                                          | .25 points                                                                                                                                                           | .50 points                                                                                                               | .75 points                                                                                                                                                                           | 1 point                                                                                                          |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|
| None of the word is written. This includes;                                                                                                                                         | 1/4 of word is written. This includes;                                                                                                                               | 1/2 of word is written. This includes;                                                                                   | 3/4 of word is written. This includes;                                                                                                                                               | Entire word is written. This includes;                                                                           |
| <ul style="list-style-type: none"> <li>• Nothing is written;</li> <li>• The letters present do not meet any 'for .25' criteria;</li> <li>• Turkish word only is written.</li> </ul> | <ul style="list-style-type: none"> <li>• Any 1 letter is correct;</li> <li>• 25-49.9% of the letters are present;</li> <li>• Correct number of syllables.</li> </ul> | <ul style="list-style-type: none"> <li>• 25-49.9% of letters correct;</li> <li>• 50-74.9% of letters present.</li> </ul> | <ul style="list-style-type: none"> <li>• 50-99.9% of letters correct;</li> <li>• 75-99.9% of letters present;</li> <li>• 100% of letters correct but other letters added.</li> </ul> | <ul style="list-style-type: none"> <li>• 100% of letters correct;</li> <li>• 100% of letters present.</li> </ul> |

**Instructions.** (1) 'Correct' refers to any letter written and placed in its correct position within a word; 'present' refers to any letter written but not placed in its correct position. (2) Determine percentages by dividing letters correct and letters present by the number of letters in the target word. If more letters are written than are in the target word, divide by the larger number. (3) If the same target is written more than once, score it only once in the space where it should be written or, if it is not written in the correct space, score it in the first space where it is written based upon the target word for that space.

**Source.** Barcroft (2000); adapted from Keating (2008).

## APPENDIX K

### NORMALITY TEST RESULTS

**Table K1. Normality Test of the Lower-Intermediate Participants**

|                                 | Groups | Statistic | <i>df</i> | <i>p</i> |
|---------------------------------|--------|-----------|-----------|----------|
| <b>Immediate<br/>Productive</b> | SR     | 0.936     | 21        | 0.184    |
|                                 | SW     | 0.961     | 18        | 0.612    |
|                                 | FR     | 0.889     | 20        | 0.025*   |
|                                 | FW     | 0.880     | 20        | 0.018*   |
|                                 | RO     | 0.783     | 21        | 0.000*   |
|                                 | WO     | 0.545     | 23        | 0.000*   |
| <b>Delayed<br/>Productive</b>   | SR     | 0.910     | 21        | 0.054    |
|                                 | SW     | 0.951     | 18        | 0.434    |
|                                 | FR     | 0.892     | 20        | 0.030*   |
|                                 | FW     | 0.627     | 20        | 0.000*   |
|                                 | RO     | 0.801     | 21        | 0.001*   |
|                                 | WO     | 0.861     | 23        | 0.004*   |
| <b>Immediate<br/>Receptive</b>  | SR     | 0.765     | 21        | 0.000*   |
|                                 | SW     | 0.838     | 18        | 0.005*   |
|                                 | FR     | 0.907     | 20        | 0.055    |
|                                 | FW     | 0.942     | 20        | 0.262    |
|                                 | RO     | 0.905     | 21        | 0.043*   |
|                                 | WO     | 0.900     | 23        | 0.025*   |
| <b>Delayed<br/>Receptive</b>    | SR     | 0.848     | 21        | 0.004*   |
|                                 | SW     | 0.770     | 18        | 0.001*   |
|                                 | FR     | 0.865     | 20        | 0.010*   |
|                                 | FW     | 0.896     | 20        | 0.034*   |
|                                 | RO     | 0.952     | 21        | 0.367    |
|                                 | WO     | 0.870     | 23        | 0.007*   |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only, \* $p < .05$

**Table K2. Normality Test of the Upper-Intermediate Participants**

|                                 | Groups | Statistic | <i>df</i> | <i>p</i> |
|---------------------------------|--------|-----------|-----------|----------|
| <b>Immediate<br/>Productive</b> | SR     | 0.943     | 19        | 0.300    |
|                                 | SW     | 0.908     | 17        | 0.091    |
|                                 | FR     | 0.959     | 17        | 0.605    |
|                                 | FW     | 0.948     | 21        | 0.306    |
|                                 | RO     | 0.524     | 19        | 0.000*   |
|                                 | WO     | 0.354     | 20        | 0.000*   |
| <b>Delayed<br/>Productive</b>   | SR     | 0.943     | 19        | 0.295    |
|                                 | SW     | 0.899     | 17        | 0.067    |
|                                 | FR     | 0.852     | 17        | 0.012*   |
|                                 | FW     | 0.901     | 21        | 0.037*   |
|                                 | RO     | 0.799     | 19        | 0.001*   |
|                                 | WO     | 0.702     | 20        | 0.000*   |
| <b>Immediate<br/>Receptive</b>  | SR     | 0.882     | 19        | 0.023*   |
|                                 | SW     | 0.817     | 17        | 0.004*   |
|                                 | FR     | 0.936     | 17        | 0.275    |
|                                 | FW     | 0.946     | 21        | 0.286    |
|                                 | RO     | 0.693     | 19        | 0.000*   |
|                                 | WO     | 0.918     | 20        | 0.091    |
| <b>Delayed<br/>Receptive</b>    | SR     | 0.926     | 19        | 0.146    |
|                                 | SW     | 0.943     | 17        | 0.352    |
|                                 | FR     | 0.925     | 17        | 0.180    |
|                                 | FW     | 0.967     | 21        | 0.660    |
|                                 | RO     | 0.202     | 19        | 0.041*   |
|                                 | WO     | 0.940     | 20        | 0.244    |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only, \* $p < .05$

## APPENDIX L

### CHANGES IN WORD KNOWLEDGE (LOWER-INTERMEDIATE LEARNERS)

**Table L1. Changes in Productive Word Knowledge (Lower-Intermediate)**

| Groups |           | <i>N</i> | $\bar{x}$ | Min. | Max. | <i>Z</i> | <i>p</i> |
|--------|-----------|----------|-----------|------|------|----------|----------|
| SR     | Immediate | 21       | 5.5       | 1.25 | 7.5  |          |          |
|        | Delayed   | 21       | 2         | 0    | 7.5  | -3.304   | 0.001*   |
| SW     | Immediate | 18       | 4         | 1    | 7.75 |          |          |
|        | Delayed   | 18       | 3.38      | 0    | 8    | -2.419   | 0.016*   |
| FR     | Immediate | 20       | 5.38      | 1    | 8    |          |          |
|        | Delayed   | 20       | 3         | 1    | 7.75 | -1.707   | 0.088    |
| FW     | Immediate | 20       | 2         | 0    | 7.75 |          |          |
|        | Delayed   | 20       | 0         | 0    | 4.5  | -3.632   | <0.001*  |
| RO     | Immediate | 21       | 0.75      | 0    | 1    |          |          |
|        | Delayed   | 21       | 0.25      | 0    | 3    | -0.930   | 0.352    |
| WO     | Immediate | 23       | 0         | 0    | 3.75 |          |          |
|        | Delayed   | 23       | 1         | 0    | 3    | -3.240   | 0.001*   |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only, \* $p < .05$

**Table L2. Changes in Receptive Word Knowledge (Lower-Intermediate)**

| Groups |           | <i>N</i> | $\bar{x}$ | Min. | Max. | <i>Z</i> | <i>p</i> |
|--------|-----------|----------|-----------|------|------|----------|----------|
| SR     | Immediate | 21       | 7         | 3    | 8    |          |          |
|        | Delayed   | 21       | 7         | 4    | 8    | -0.597   | 0.551    |
| SW     | Immediate | 18       | 7         | 3    | 8    |          |          |
|        | Delayed   | 18       | 7.5       | 2    | 8    | -0.071   | 0.943    |
| FR     | Immediate | 20       | 6.25      | 2    | 8    |          |          |
|        | Delayed   | 20       | 6.25      | 0    | 8    | -1.064   | 0.287    |
| FW     | Immediate | 20       | 3.75      | 0    | 8    |          |          |
|        | Delayed   | 20       | 2         | 0    | 8    | -2.437   | 0.015*   |
| RO     | Immediate | 21       | 2         | 0    | 6    |          |          |
|        | Delayed   | 21       | 2         | 0    | 6    | -0.087   | 0.930    |
| WO     | Immediate | 23       | 4         | 0    | 6    |          |          |
|        | Delayed   | 23       | 3         | 1    | 6    | -0.105   | 0.916    |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only, \* $p < .05$

## APPENDIX M

### CHANGES IN WORD KNOWLEDGE (UPPER-INTERMEDIATE LEARNERS)

**Table M1. Changes in Productive Word Knowledge (Upper-Intermediate)**

| Groups |           | <i>N</i> | $\bar{x}$ | Min. | Max. | <i>Z</i> | <i>p</i> |
|--------|-----------|----------|-----------|------|------|----------|----------|
| SR     | Immediate | 19       | 3.75      | 0    | 8    |          |          |
|        | Delayed   | 19       | 2         | 0    | 4.5  | -2.702   | 0.007*   |
| SW     | Immediate | 17       | 4.75      | 0.25 | 8    |          |          |
|        | Delayed   | 17       | 2.75      | 0    | 6    | -3.623   | <0.001*  |
| FR     | Immediate | 17       | 3         | 0    | 7.5  |          |          |
|        | Delayed   | 17       | 1.75      | 0    | 3    | -3.424   | 0.001*   |
| FW     | Immediate | 21       | 4.75      | 0    | 8    |          |          |
|        | Delayed   | 21       | 1.5       | 0    | 7    | -2.614   | 0.009*   |
| RO     | Immediate | 19       | 0         | 0    | 3    |          |          |
|        | Delayed   | 19       | 1         | 0    | 3    | -2.470   | 0.014*   |
| WO     | Immediate | 20       | 0         | 0    | 2    |          |          |
|        | Delayed   | 20       | 1         | 0    | 4    | -2.812   | 0.005*   |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only, \* $p < .05$

**Table M2. Changes in Receptive Word Knowledge (Upper-Intermediate)**

| Groups |           | <i>N</i> | $\bar{x}$ | Min. | Max. | <i>Z</i> | <i>p</i> |
|--------|-----------|----------|-----------|------|------|----------|----------|
| SR     | Immediate | 19       | 6         | 1    | 8    |          |          |
|        | Delayed   | 19       | 5         | 1.5  | 8    | -0.641   | 0.521    |
| SW     | Immediate | 17       | 7         | 3    | 8    |          |          |
|        | Delayed   | 17       | 5         | 1    | 7    | -3.319   | 0.001*   |
| FR     | Immediate | 17       | 4         | 1    | 7    |          |          |
|        | Delayed   | 17       | 3         | 0    | 6    | -2.466   | 0.014*   |
| FW     | Immediate | 21       | 6         | 1    | 8    |          |          |
|        | Delayed   | 21       | 3.5       | 0    | 6.5  | -2.473   | 0.013*   |
| RO     | Immediate | 19       | 1         | 0    | 6    |          |          |
|        | Delayed   | 19       | 2         | 0    | 5    | -1.696   | 0.090    |
| WO     | Immediate | 20       | 2         | 0    | 5    |          |          |
|        | Delayed   | 20       | 2.5       | 0    | 7    | -1.771   | 0.077    |

**Note.** SR: Sentence-writing-by-reading, SW: Sentence-writing-by-watching, FR: Fill-in-by-reading, FW: Fill-in-by-watching, RO: Reading-only, WO: Watching-only, \* $p < .05$