T.C.

BURSA ULUDAG UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

FOREIGN LANGUAGE EDUCATION

DEPARTMENT OF ENGLISH LANGUAGE TEACHING

THE IMPACT OF MACHINE TRANSLATION ON CREATIVITY IN WRITING TASKS IN EFL CLASSES

MASTER THESIS

Ayşe TUZCU

BURSA

2020

T.C.

BURSA ULUDAG UNIVERSITY

INSTITUTE OF EDUCATIOANAL SCIENCES

FOREIGN LANGUAGE EDUCATION

DEPARTMENT OF ENGLISH LANGUAGE TEACHING

THE IMPACT OF MACHINE TRANSLATION ON CREATIVITY IN WRITING TASKS IN EFL CLASSES

MASTER THESIS

Ayşe TUZCU

Supervisor

Prof. Dr. Zübeyde Sinem GENÇ

BURSA

2020

BİLİMSEL ETİĞE UYGUNLUK

Bu çalışmadaki tüm bilgilerin akademik ve etik kurallara uygun bir şekilde elde edildiğini beyan ederim.

Ayşe TUZCU

04 / 02 / 2020

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS İNTİHAL YAZILIM RAPORU

ULUDAĞ ÜNİVERSİTESİ EĞİTİM BİLİMLER ENSTİTÜSÜ

YABANCI DİLLER EĞİTİMİ ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 02/03/2020

Tez Başlığı / Konusu: Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi

Yukarıda başlığı gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 105 sayfalık kısmına ilişkin, 02/03/2020 tarihinde şahsım tarafından (Turnitin)* adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan özgünlük raporuna göre, tezimin benzerlik oranı % 5'tir.

Uygulanan filtrelemeler:

1- Kaynakça hariç

2- Alıntılar hariç

3- 5 kelimeden daha az örtüşme içeren metin kısımları hariç

Uludağ Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Özgünlük Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Adı Soyadı : Ayşe TUZCU

02 / 03 / 2020

Ayşe TUZCU

Öğrenci nu : 801610021

Anabilim Dalı : Yabancı Diller Eğitimi

Anabilim Dalı Programı : İngilizce Öğretmenliği

Programı Statüsü : Yüksek Lisans

Prof. Dr. Zübeyde Sinem GENÇ

03/03/2020

* Turnitin programına Uludağ Üniversitesi Kütüphane web sayfasından ulaşılabilir.

YÖNERGEYE UYGUNLUK ONAYI

"Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi" adlı Yüksek Lisans Tezi, Uludağ Üniversitesi Eğitim Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırlanmıştır.

Tezi Hazırlayan

Ayşe TUZCU

Danışman

Prof. Dr. Zübeyde Sinem GENÇ

Yabancı Diller Eğitimi Anabilim Dalı Başkanı

Prof. Dr. Zübeyde Sinem GENÇ

T.C.

BURSA ULUDAĞ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Eğitim Bilimleri Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Bilim Dalı öğrencisi 801610021 numara ile Ayşe TUZCU'nun hazırladığı "Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi" konulu Yüksek Lisans çalışması ile ilgili tez savunma sınavı, 08 / 05 /2020 günü 14.30 – 15.30 saatleri arasında yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin/çalışmasının **başarılı** olduğuna **oybirliği** ile karar verilmiştir.

Üye (Tez Danışmanı ve Sınav Komisyonu Başkanı)

Prof. Dr. Zübeyde Sinem GENÇ

Üye

Dr Öğr. Üyesi Kürşat CESUR

Üye

Dr Öğr. Üyesi Derya YILMAZ

Abstract

Author : Ayşe TUZCU

University : T.C. Bursa Uludag University

Field : Foreign Languages Education

Branch : English Language

Degree Awarded : MA Thesis

Page Number : xix+ 159

Degree Date :

Thesis : The Impact of Machine Translation on Creativity in Writing Tasks in EFL Classes

Supervisor : Prof. Dr. Zübeyde Sinem GENÇ

THE IMPACT OF MACHINE TRANSLATION ON CREATIVITY IN WRITING TASKS IN EFL CLASSES

Due to the fact that there is a limited number of research on machine translation as a learning tool, it is still controversial whether it is effective or not. The aspect of those studies is also narrow. Most research approaches machine translation in terms of its impact on grammar and lexis or the perceptions of the users and the strategies they generated. However, a very important point is neglected: creativity.

The concern of how to flourish creativity in EFL classes has long been investigated in the field of education. The current study combines the literature of these two different subjects – machine translation and creativity- and scrutinizes the impact of machine translation, Google Translate in this research, on creativity of beginner learners in writing tasks in EFL classes.

It was conducted in a state Anatolian high school in Bursa, Turkey between September 2, 2019 and February 29, 2020 in the first semester of 2019-2020 educational year. It had a one-group quasi-experimental design with 35 beginner EFL (English as a foreign language) voluntary learners, whose parents had given a consent form. The study lasted for ten weeks including a training session, pre-and post-tests. Quantitative data were collected through a pre and post-test design and were analysed in terms of both creativity, grammatical and lexical errors with paired samples t-test and Wilcoxon test. The inter-rater coefficient alphas in the CAT analysis are found to be .86 for the pre-test and .85 for the post-test. In addition, the differences between the fluency, flexibility, originality and elaboration scores between the pre and post-test are statistically significant while no

differences are detected between the grammatical and lexical error scores. Qualitative data were gathered through a semi-structured interview and the responses show that the students have a positive perception of implementing Google Translate in their writing classes.

Key words: the Consensual Assessment Technique (the CAT), creativity, machine translation, writing skills

Özet

Yazar : Ayşe TUZCU

Üniversite : T. C. Bursa Uludağ Üniversitesi

Ana Bilim Dalı : Yabancı Diller Eğitimi Anabilim Dalı

Bilim Dalı : İngiliz Dili Eğitimi Bilim Dalı

Tezin Niteliği : Yüksek Lisans Tezi

Sayfa Sayısı : xix + 159

Mezuniyet Tarihi :

Tez : Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi

Danışmanı : Prof. Dr. Zübeyde Sinem GENÇ

YABANCI DİL OLARAK İNGİLİZCE ÖĞRENİLEN SINIFLARDA MAKİNE ÇEVİRİSİNİN YAZMA GÖREVLERİNDE YARATICILIĞA ETKİSİ

Son yıllarda makine çevirisi teknolojisini bir eğitim aracı olarak gören bazı çalışmalar yapıldı. Ne var ki, bu çalışmaların az sayıda olmalarından dolayı makine çevirisinin etkili bir araç olup olmadığı konusunda henüz bir anlaşmaya varılamadı. Bu çalışmaların sınırlı sayıda olması bir yana konuya dair bakış açıları da sınırlı. Birçok araştırma, makine çevirisinin dil ve kelime bilgisi üzerine etkisini, kullanıcıların algılarını ya da kullanırken başvurdukları stratejileri belirlemeye yönelik incelemeler yaptı. Ne yazık ki çok önemli bir nokta göz ardı edildi: yaratıcılık.

Yaratıcılık eğitim alanında yeni okur-yazarlık becerileri arasında önemli bir role sahiptir. Yabancı dil olarak İngilizce öğretilen sınıflarda yaratıcılığın nasıl arttırılıp geliştirileceği kaygısı şimdiye kadar birçok farklı bakış açısıyla incelenmiştir. Bu çalışma bu iki farklı literatürü – makine çevirisi ve yaratıcılık – bir araya getirerek makine çevirisi, bu çalışmada Google Çeviri, kullanımının başlangıç düzeyinde İngilizce bilen lise öğrencilerinin yazma etkinliklerinde yaratıcılığına olan etkisini incelemeyi hedeflemektedir.

Bu çalışma, Türkiye'nin Bursa ilinde T.C. Milli Eğitim Bakanlığı'na bağlı bir Anadolu lisesinde 2019-2020 eğitim – öğretim yılının ilk yarısında 02.09.2019 ve 29.02.2020 tarihleri arasında yürütüldü. Tek gruplu yarı-deneysel bir çalışma olan araştırmada başlangıç düzeyinde İngilizce bilen, veli onam formları alınmış ve kendileri de gönüllü 35 dokuzuncu sınıf öğrencisiyle çalışıldı. Uygulama, ön-test ve son-testler de dâhil olmak üzere on hafta sürdü. Nicel veri ön ve son testler sırasında yazılan metinlerin

incelenmesiyle elde edildi. Metinler hem yaratıcılık hem de dilbilgisi ve kelime bilgisi açısından eşleştirilmiş örneklem t-test ve Wilcoxon testiyle analiz edildi. Uzlaşmacı Değerlendirme Tekniği analizinde değerlendiriciler arası güvenirlik katsayısı ön-test için .86, son-test için .85 olarak tespit edilmiştir. Buna ek olarak, ön ve son testlerdeki akıcılık, esneklik, özgünlük ve detaylandırma skorları arasında istatistiksel olarak anlamlı farklar oluşmuşsa da dilbilgisi ve kelime hatalarında testler arasında bir fark gözlenmemiştir. Yarı yapılandırılmış görüşmelerden elde edilen nicel veri sonuçları katılımcıların yazma etkinliklerinde Google Çeviri kullanmaya dair olumlu algılara sahip olduklarını göstermiştir.

Anahtar sözcükler: makine çevirisi, Uzlaşmacı Değerlendirme Tekniği, yaratıcılık, yazma becerileri

Acknowledgements

I would like to present my appreciation to those who are standing beside me through this demanding yet amazing process. First of all, I am extremely grateful to my supervisor, Prof. Dr. Zübeyde Sinem Genç, with her priceless guidance. Without her scientific point of view, this study would not be as fruitful as it is now. This challenging process was more productive and more pleasant with her guidance. I will always be thankful to my supervisor for everything I learnt from her in order to develop as a better teacher.

I am also grateful to Assoc. Prof. Dr. Deniz Sığırlı for her contribution to the study. She provided an invaluable assistance with the analysis of the data set.

I want to express my gratitude to my dear colleagues who took part voluntarily in the creativity analysis of the texts in terms of the CAT without a little sign of hesitance. Without their precious assistance, this study would never reach its aim. In addition to my colleagues, my dear students participated in the study voluntarily and showed a great patience during the every step of the study. They were the real heroes and heroines of this journey.

Last but not the least, the special thanks go to my lovely family who always stand beside me. I dedicate this work to my little cute two-year old son, Tolga Tuzcu and I hope I can be an idol in his life and help him appreciate how invaluable every single piece of scientific knowledge is.

Ayşe TUZCU

Table of Contents

Abstract vi

Özet viii

Acknowledgements x

Table of Contents xi

List of Tables xvii

List of Figures xviii

List of Abbreviations xix

Chapter 1 1

Introduction 1

1.1 The Background of the Study 1

1.2 The Statement of the Problem 2

1.3. Purpose of the Study 3

1.4. Research Questions 4

1.5. Significance of the Study 4

1.6. Limitations 5

1.7. Definitions of the Terms 6

Chapter 2 7

Literature Review 7

2.1. Blended Learning 7

2.1.1. Machine translation. 9

2.1.1.1. Strengths of machine translation. 10

2.1.1.2. Weaknesses of machine translation. 11

2.1.2. Foreign language learning and machine translation. 12
2.2 Creativity 17
2.2.1. Definition. 17
2.2.2. Creativity and Joy Paul Guilford. 20
2.2.2.1. Fluency 22
2.2.2.2. Flexibility 23
2.2.2.3. Originality. 24
2.2.2.4. Elaboration 24
2.2.3. Creativity and language 25
2.2.4. Creativity and education 26
2.2.5. Measurement and assessment 29
2.2.5.1. Reliability and validity of creativity tests. 31
2.2.5.2. The Consensual Assessment Technique 33
2.3. Surface Structure Taxonomy 37
2.3.1. Omission. 38
2.3.2. Addition 38
2.3.3. Misformation. 39
2.3.4. Misordering 39
Chapter 3 41
Methodology 41
3.1. Introduction 41
3.2. Aim of the Study and Research Questions 41

3.3. Research Design 41

3.4. Context 45

3.5. Participants 46

3.5.1. The students. 46

3.5.2. Role of the researcher. 47

3.5.3. The raters 47

3.6. Data Collection Instruments 47

3.6.1. Google Translate 47

3.6.2. Semi-structured interview. 48

3.7. Data Collection Process 48

3.7.1. The implementation of the pre-test and the post-test. 50

3.7.2. The instruction process 52

3.7.2.1. Training session 53

3.7.2.2. Week 1 54

3.7.2.3. Week 2 55

3.7.2.4. Week 3 56

3.7.2.5. Week 4 57

3.7.2.6. Week 5 58

3.7.2.7. Week 6 59

3.7.2.8 Week 7 60

3.8. Data Analysis Procedure 61

3.8.1. Creativity assessment 61

3.8.2. Grammatical and lexical error analysis 64

Chapter 4 65

Results 65

4.1. Introduction 65

4.2. Results of the Quantitative Data Set 65

4.2.1. Creativity 65

4.2.1.1 The Consensual Assessment Technique 65

4.2.1.2. Components of creativity in pre and post-tests. 67

4.2.1.2.1. Fluency in the pre and post-tests. 68

4.2.1.2.2. Flexibility in the pre and post-tests. 68

4.2.1.2.3. Originality in the pre and post-tests. 69

4.2.1.2.4. Elaboration in the pre and post-tests 70

4.2.2. Grammatical and lexical knowledge 71

4.2.2.1. Omission errors 71

4.2.2.2. Addition errors. 72

4.2.2.3. Misformation errors 72

4.2.2.4. Misordering. 73

4.2.2.5. Lexical choice errors. 74

4.3. Results of the Qualitative Data Set 74

4.3.1. Question 1: What were the advantages and disadvantages of implementing machine translation in writing activities?. 75

4.3.2. Question 2: Do you think implementing machine translation affected your creativity in writing activities? 79
4.3.3. Question 3: Do you feel like you learnt new things about English while implementing machine translation in writing activities?. 81
4.3.4. Question 4: Did you experienced any problems while using machine translation?. 84
4.3.5. Question 5: Had you used machine translation before we started this study?. 84
Chapter 5 87
Discussion 87
5.1. Introduction 87
5.2. Creativity in the Written Products 88
5.2.1. The Consensual Assessment Technique. 88
5.2.2. Components of creativity in pre and post-tests 93
5.3. Grammatical and Lexical Knowledge 96
5.4. The Perceptions of Participants 98
5.5. Summary 100
Chapter 6 101
Conclusion 101
6.1. Implications of the Study for EFL Teachers 103
6.2. Suggestions for Further Research 104
References 106
Appendices 125
Appendix 1: Background Information Form 125

Appendix 2: Samples from the Products in the Pre-test 126

Appendix 3: The Prompt Image for the Post-test 130

Appendix 4: Samples from the Products in the Post-test 131

Appendix 5: Traning Session 135

Appendix 6: Samples from the First Week Activity 139

Appendix 7: Samples from the Second Week Activity 141

Appendix 8: Samples from the Third Week Activity 143

Appendix 9: Samples from the Forth Week Activity 145

Appendix 10: Samples from the Fifth Week Activity 146

Appendix 11: Samples from the Sixth Week 148

Appendix 12: Samples from the Seventh Week Activities 150

Appendix 13: Analysis of Grammatical and Lexical Errors in Pre and Post-tests 152

Appendix 14: Samples from Mistranslation on Google Translate 154

Appendix 15: Bursa Provincial Directorate of National Education Approval Letter 155

Appendix 16: Uludağ University Research Ethics Committee Approval Letter 156

Appendix 17: Institute of Educational Sciences Committee Approval 157

List of Tables

Table *Page*

1. The data collection process ……. 49
2. Descriptive statistics of the raters for pre and post-tests …….. 67
3. Descriptive statistics for the fluency scores in pre and post-tests ……... 68
4. Descriptive statistics for the flexibility scores in pre and post-tests …… 69
5. Descriptive statistics for the originality scores in pre and post-tests ……... 70
6. Descriptive statistics for the elaboration scores in pre and post-tests …….. 70
7. Descriptive statistics for the omission errors in pre and post-tests ……... 72
8. Descriptive statistics for the addition errors in pre and post-tests …… 72
9. Descriptive statistics for the misformation errors in pre and post-tests …… 73
10. Descriptive statistics for the misordering errors in pre and post-tests …….. 73
11. Descriptive statistics for the lexical choice errors in pre and post-tests …… 74
12. Summary of the categories, subcategories and the corresponding codes emerging from the first question in the semi-structured interviews ……..... 77
13. Summary of the categories, subcategories and the corresponding codes emerging from the second question in the semi-structured interviews ……. 80
14. Summary of the categories, subcategories and the corresponding codes emerging from the third question in the semi-structured interviews ……..... 82
15. Summary of the categories, subcategories and the corresponding codes emerging from the forth question in the semi-structured interviews ……... 84
16. Summary of the categories, subcategories and the corresponding codes emerging from the fifth question in the semi-structured interviews ……..... 86

List of Figures

Figure *Page*

1. Mistranslation of an idiom on https://translate.google.com ……... 12

2. Mistranslation of a poem by Shakespeare on https://translate.google.com …… 12

3. Mistranslation due to inaccurate spelling on https://translate.google.com …… 54

4. Output due to accurate spelling on https://translate.google.com ……..... 54

List of Abbreviations

ATC: Abedi Test of Creativity

ATTA: Abbreviated Torrance Test for Adults

CALL: Computer Assisted Language Learning

CAT: Consensual Assessment Technique

CEFR: Common European Framework of Reference for Languages

DYNED: Dynamic Education

EFL: English as a Foreign Language

GT: Google Translate

ICC: Intra-class Correlation Coefficient

L1: First Language

L2: Second Language

LAP: Languages for Academic Purposes

MALL: Mobile Assisted Language Learning

MoNE: Ministry of National Education

MT: Machine Translation

SOI: Structure-of-Intellect

SPSS: Statistical Package for Social Sciences

TTCT: Torrance Test of Creative Thinking

Chapter 1

Introduction

The present study consists of six chapters. The first chapter is the introduction part and explains the background of the study, the statement of the problem, the purpose of the study, research questions and the significance of the study. The section is concluded with the definitions of the terms.

1.1 The Background of the Study

Machine translation (MT) has been increasingly used by foreign language learners especially by the ones with low level of proficiency as an assistant for their learning tasks. A recent study by Alhaisoni and Alhaysony (2017) suggests that learners prefer addressing MT for vocabulary, reading and writing. On the other hand, it is shown that the least frequently mentioned reason is translation (Alhaisoni & Alhaysony, 2017). Similarly, in a study by White and Heidrich (White & Heidrich, 2013) it is revealed that most of the participants are eager to use MT in writing activities even though they are not allowed to use mobile phones in the classroom.

To date, there has been a number of research on MT from many different points of views; however, there is little study about its impact on creativity in writing tasks, although creativity has been a top topic for many studies almost for more than a century. As J. P. Guilford, one of the pioneers in the field of creativity, suggests, creativity should be considered under two titles: “creative potential” and “creative performance” (Guilford, 1966, p. 186). According to him, creative potential is what a person actually possesses and makes use of when creativity needed; on the other hand, creative performance is described as what a person actually produces with the help of his/her potential and the other situational conditions. Considering this information, it can be inferred that in writing tasks where EFL learners should turn their creative potential into creative performance they have difficulty in putting

their ideas into words especially the ones with low level of proficiency. Henceforth, they need an assistant: why not this assistant be a machine translation tool?

1.2 The Statement of the Problem

Even though there is still a debate on the use of machine translation in EFL classes, it is obvious in previous studies that students are inclined to make use of it during their language learning process. Among the limited number of studies that focus on the role of MT as a pedagogical tool, some of them reveal the fact that MT can be an appropriate learning tool in EFL classes. For example, a study by Sangmin-Michelle Lee reveals that the participants are in favour of using MT in their writing; beside, MT has also a positive effect on decreasing their lexico-grammatical errors in writing (Lee, 2019). Another research which proves that students use MT in learning a different language is held by Kumar A. (2012). According to this research, the students use MT for many reasons: understanding questions, topics, concepts (Kumar A., 2012).

Although researchers have been trying to find out efficient ways of implementing that technological tool which is already used by EFL students either in-class or out-class activities, its impact on creativity is being neglected. Creativity is considered as one of the most popular requirements of the new era and although it has been defined by a great number of researchers in a variety of ways, a definition by Terese Amabile (2012, p.3) explains it as “the production of a novel and appropriate response, product, or solution to an open-ended task”. As it is obvious, the core of the definition is a novel and appropriate production, which straightly addresses to the ratio behind the writing skill in EFL classes, where students are supposed to create novel and appropriate pieces of writings in line with the given instructions. However, especially the beginner learners have problems with their lack of grammatical and lexical knowledge and they need scaffolding. Then, can machine translation impact the students’ ability to produce novel and appropriate solutions in their writing activities? In other words,

can the low-proficient EFL learners create novel and appropriate written products with the help of a machine translation and write more grammatically accurate sentences?

In the present research, another concern about the machine translation is the perceptions of the students. Aside from the impact of the machine translation on the creativity and the accuracy in grammar and lexis, what the learners think about utilizing a machine translator in their writing activities in English classes is another essential issue that will be investigated in the study.

1.3. Purpose of the Study

Acknowledging the fact that new generation is born into the technology, it is essential to investigate how to implement technological devices in language learning environments by discovering both the potential benefits and the danger they pose. In order to reach this aim, the present study investigates about the impact of the Google Translate (https://translate.google.com.tr/), as a machine translator, on creativity, which is a requirement of the new era, along with the grammatical and lexical competence of EFL students with low level of English proficiency in an Anatolian high school in Bursa. It is expected to fill this gap in the field and to propose sufficient questions and answers for further studies.

In other words, the main aim of the present study is to present an insight into the impact of using a machine translator the creativity in written products of low proficient EFL learners. In order to do this, a two-phased assessment process is administered. The first one is the Consensual Assessment Technique, which was first introduced by Amabile (CAT; Amabile, 1982). The technique makes it possible for raters who are experts in the same domain as the products to rate subjectively the products relative to each other. A rank-order of the written products in the pre and post-tests in terms of creativity is aimed at by means of the CAT.

Moreover, the four components of creativity, suggested by Guilford (1966, p. 188), as flexibility, fluency, originality and elaboration, are taken into consideration and will be evaluated separately both in the pre and post-tests in order to see if there is any significant difference in between.

In addition to creativity, this paper aims at defining the impact of MT on the students' knowledge of grammar and lexis knowledge by evaluating the pre and post-tests in terms of grammatical errors in line with the Surface Strategy Taxonomy (Dulay, Burt & Krashen, 1982, p. 150 - 163) and in terms of lexical errors in line with "lexical choice and word form" (Barkaoui, K., & Knouzi, I., 2012, p. 93). Lastly, the perceptions of participants on the use of MT as a learning tool will be investigated by the help of semi-structured interviews.

1.4. Research Questions

The research questions of this study are as follows:

RQ1. How does implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, elaboration and originality?

RQ2. How does implementing machine translation in writing activities in English affect the participants' grammatical and lexical knowledge?

RQ3. What are the perceptions of high school EFL students on the use of machine translation in writing activities in English classes?

1.5. Significance of the Study

Although the utilisation of translation in foreign language teaching and learning process was despised for many reasons, it has long been appreciated by some scientific research that proves its usefulness as a beneficial activity to assist foreign language learning (Fernández-Guerra, 2014). According to Leonardi (2010, p. 23) translation is not just a mechanical activity as it was labelled beforehand, it is rather a complex process that includes four skills - reading, writing, listening and speaking.

As translation is gaining its reputation again as a teaching and learning activity, it has been accompanied with another feature: technology. Technological developments introduced the users with machine translation in a number of languages. This technology is accessible for many of our students in or outside of our EFL classes. However, there is a limited number of studies that focuses on the potential effects of using machine translation. While learners with low proficiency levels do inevitably address to the machine translation for any of their tasks, we, as teachers, should first consider about its potentials before we ban its use. Thus, the study is expected to shed light on one of the effective ways of using machine translation with beginner learners in writing classes for EFL teachers.

Creativity in EFL learning environments, which is one of the main elements in the present research, has long been studied by many researchers (see Dikici, 2013; Gursoy & Bag, 2018; Henriksen, 2016; Kırkgöz, 2014; Pishghadam & Mehr, 2011). Nevertheless, its relationship with a technological tool, machine translation in this research, especially with low proficient high school students, has not been in the spotlight of much research.

Taking these into consideration, this research is expected to fill in this gap as being significant in that it may find answers to questions that are not frequently asked in the previous studies in education. The present study with its two-phased creativity assessment, beginner level of participants and the combination of creativity and the machine translation is expected to be significant in the educational research field.

1.6. Limitations

This study was conducted as a quasi-experimental research with a single group. As listed by Thompson and Panacek (2006) one of the most frequent quasi-experimental design is the group sequential design, which is called as single group time series. They explain the process briefly as “A single population of subjects is selected and used as its own controls as it goes through a series of observations and interventions, all in the same order” (Thompson &

Panacek, 2006, p. 245). As the study was conducted in a state high school, although the participants could not be assigned randomly, one of the seven 9th grade classes was chosen randomly with a class size of thirty-five students.

Additionally, the study lacks a control group. In educational research, it is suggested by many researchers to have a control group (Dörnyei, 2011); however, in order to be able to control the possibility of unexpected variables, such as the level of the learners, the impact of the instructor, the control group is not included in the study.

The number of the participants is 35, which is too small to generalise the results. That is because it is hard to reach more numbers of participants in a single classroom. Nevertheless, the study is expected to trigger a line of research on machine translation and its impact on creativity shedding a light on the need of research and a promising way to inquire.

1.7. Definitions of the Terms

These terms are defined in line with the aim of this present study:

Machine translation: "Machine translation (MT) is the application of computers to the task of translating texts from one natural language to another"(Okpor, 2014, p. 159).

Creativity: "The production of a novel and appropriate response, product, or solution to an open-ended task" (Amabile, 2012, p. 3).

Chapter 2

Literature Review

This chapter gives a detailed insight about the ratio behind the study. As being the core subjects of the study, the literature about machine translation and creativity is probed into under several subtitles. The idea behind machine translation, the strengths and weaknesses as a learning tool and the previous research are described in detail. Secondly, creativity as a multi-layered subject, is explained in terms of its definition, its pioneers, components, the assessment techniques and validity and reliability in creativity research. Lastly, as a secondary subject of this research, surface structure taxonomy is explained with suitable examples for every topic it offers.

2.1. Blended Learning

Technology is inevitably in every field of our life. Thanks to the rapid developments in technology, there have been numerous changes in the way that people exchange information, facing less problems about time limitation and accessibility (Özerbaş and Mart, 2017). These dramatic changes naturally affect the course of education as it does in other fields like medicine, business, and social sciences and thus increasing number of studies that focus on the technology integration into education has come up with a term: blended learning.

Blended learning, a relatively new term in the field of education, has been used since at least 2007 (Hockly, 2018). Sharrma and Barrett published their eponymous book and first introduced the term to the field. As Tomlinson and Whittaker (2013) mentions in their study on blended learning, the term was first used to refer to the technology use in the business world, then was addressed in higher education and finally appeared in the language teaching and learning stages.

The most generally referred definition of blended learning in language learning and teaching goes as any combination of face-to-face teaching with computer technology (online

and offline activities/materials) (Tomlinson & Whittaker, 2013). As is it clear from the definition that blended learning does not exclude face to face teaching-learning environment while adapting technology into education. Furthermore, a harmonious balance between these two elements is suggested by the researchers in this field.

Among the ways utilised in blending a course, Mobile-Assisted Language Learning (MALL) is by far the most popular technique, which involves using mobile devices like iPods, iPads, mobile phones, tablets, GPS tools, laptop computers, MP3 or MP4 players, videotapes and multimedia players (Ekinci & Ekinci, 2017). Smart phones are the most favourite ones for the fact that almost every single person has a smart phone makes it easier to bring technology into EFL classes even if the school lacks in technological supplements. What is more, studies show that both learners and teachers find mobile devices very useful and necessary in their learning and teaching experience (Ozer & Kilic, 2015) and they like the idea of getting instant and easy access to information via mobile devices whenever and wherever they want (Yurdagül & Oz, 2018).

Mobile devices, especially smart phones, offer learners a great number of opportunities with the help of mobile applications. In their case study Ekinci E. and Ekinci M. (2017) question the perceptions of 20 EFL learners of a state university about using mobile applications (Duolingo, Memrise, Acobot, VoScreen and English Central) for English language learning. The results show that the participants have positive perceptions of these applications in that they help learners to learn new vocabulary along with enhancing their reading, writing and listening skills; they are portable, authentic and motivating; however, they state some negative perceptions of the micropayment of some concepts and necessity to have an Internet access.

Nevertheless, Bartholomew and Reeve (2018) list the drawbacks of using mobile devices in education in nine items reviewing the previous research. The most frequently

mentioned drawbacks are that mobile devices distract students' attention, increase disciplinary problems, decrease student engagement and raise privacy concerns among students. Additionally, the administrative restriction on the mobile device use in classrooms stands as a major problem in most of the states. The findings of their own experimental study among K12 learners also suggest that mobile devices may not have a positive influence on middle school students.

In order to overcome such practical cons Hockly (2018) suggests some principles of blended learning implementation in EFL teaching and learning process. She suggests interaction among the learners, the teacher and possibly with other learners out of the classroom via tasks that are designed to meet needs in accordance with the aim of the selected technological tool. The blending process should also reflect the SLA research findings and the material should not only provide input but also facilitate the process of language learning. Another suggested area is the integration process which should include the face to face communication and technology supporting each other. Finally, she suggests a successful teacher and learner trainings to help them get the scope of the issue.

Taking all of these issues into consideration, the present study is concerned with the use of an online application in writing activities in EFL classes as a tool which is a machine translator: Google Translate and search for its impact on creativity and the grammar and lexis knowledge of the participants along with their perceptions on utilizing Google Translate.

2.1.1. Machine translation. The idea of machine translation stems from the ideal of creating a universal language (Rehm, Sasaki, Stein & Witt, 2018). It can simply be defined as "the process by which computer software is used to translate and compatible with PC systems and smart phones" (Lee, 2019). Starting as a mechanical bilingual dictionary in France in 1930s, this system emerged into a wildly used tool with the help of vast growing computer and the Internet technology after Warren Weaver, who is one of the pioneers in machine

translation, first came with the idea of using computers to make translation in 1947 in a letter to Professor Norbert Wiener of Massachusetts Institute of Technology:

> One thing I wanted to ask you about is this. A most serious problem, for UNESCO and for the constructive and peaceful future of the planet, is the problem of translation, as it unavoidably affects the communication between peoples. Huxley has recently told me that they are appalled by the magnitude and the importance of the translation job. Recognizing fully, even though necessarily vaguely, the semantic difficulties because of multiple meanings, etc., I have wondered if it were unthinkable to design a computer which would translate. Even if it would translate only scientific material (where the semantic difficulties are very notably less), and even if it did produce an inelegant (but intelligible) result, it would seem to me worthwhile (Weaver, 1949, p. 4).

Since then, machine translation has developed incredibly with the help of numerous research. Today it provides translation in a number of languages in addition to being a promising pedagogical tool for language learners.

2.1.1.1. Strengths of machine translation. There are a number of online machine translators which are widely used around the word; however, as being the main instrument in this present research, Google Translate is specifically in the spot in this subtitle.

First of all, Google Translate provides a website which offers a guideline for every language on Google Translate (visit https://translate.google.com/intl/en/about/languages/). It offers nine features for English to Turkish and backward translations: type, write, talk, speak, listen, snap, see, off-line and instant web-page translation. The ‘type’ feature lets the user type a single word or a text and get a translation. The ‘write’ feature lets the user draw letters on the screen rather than typing them. The ‘talk’ mode helps to have bilingual conversations. The ‘speak’ mode is to say a word or phrase and get the translation. The ‘listen’ feature provides

the pronunciation of the translation and lastly the ‘snap’ mode helps the users to get a translation by simply taking a snap of the source text via the camera of a smart device. What is more you can get off-line translation once you download the languages.

Secondly, Google Translate is good at conjugating basic enough verbs for beginner or pre-intermediate L2 learners to get so successful translations as to give a hint that the students use the Google Translate (Ducar & Schocket, 2018). For example, the expression ‘var’ in Turkish can be translated into English in two different forms: there is /are and have / has. The beginner level students can sometimes have problems to pick the correct equivalent. However, Google Translate can get the difference in basic level sentences. While the sentence ‘Evde yeşil bir halı var’ is translated into English as ‘There is a green carpet in the house’, ‘Yeşil bir halım var’ is translated correctly as ‘I have a green carpet’ (The example sentences were retrieved from https://translate.google.com on November 23, 2019).

Lastly, the Google formed an online translate community among volunteer Internet users in order for them to contribute to Google Translate by means of translating phrases, which means that the translations are verified by human beings as much as possible.

2.1.1.2. Weaknesses of machine translation. Google Translation is an improving online tool. Nevertheless, it sometimes fails in translation and produces errors. The most typical of these errors result from misspelling or punctuation errors. The users should type machine-translation-friendly sentences in order to get successful translations (Correa, 2014). Google Translate cannot translate the metaphorical sentences and some of the idioms. For example, if you type ‘Keçileri kaçırdı’, you get ‘He missed the goats’ (Retrieved from https://translate.google.com on November 23, 2019).

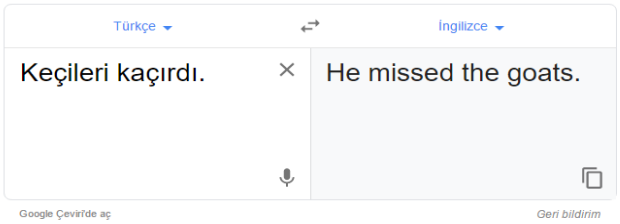


Figure 1. Mistranslation of an idiom on https://translate.google.com

One of the most common failures of Google Translate is about intercultural issues. For instance, it cannot translate a poem (Retrieved from https://translate.google.com on November, 23, 2019):

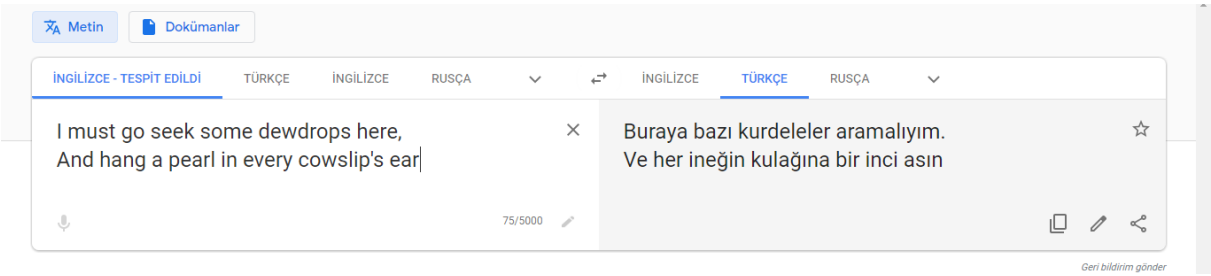


Figure 2. Mistranslation of a poem by Shakespeare on https://translate.google.com

Although a poem is a complex piece of text, it sometimes fails in determining the level of formality of simple expressions. Greeting someone especially is a very basic issue in a language but you should be careful about whom you are greeting. If one sounds informal to a future boss in a first meeting it may not be a proper introduction. When you consult Google Translate about ‘Nasılsınız?’ it translates it as ‘How are you doing?’; however, it does not mention about the level of formality (Retrieved from https://translate.google.com on November, 23, 2019).

2.1.2. Foreign language learning and machine translation. The role of translation in foreign language learning has long been subject to many discussions. The Grammar Translation method, in particular, was criticised of neglecting communication aspect of language and it was replaced by communicative methodologies; however, translation has

gained its reputation again and now is regarded as a way of communication (Saroukhil, Ghalkhani & Hashemi, 2018).

Throughout the history, translation had undergone a number of changes until technology unavoidably affected the way, the speed and the frequency that people make translation. Hence, machine translation altered translation and has now become an essential part of communication in this global world. Accordingly, with the growing technology of smart phones, the accessibility of machine translation has dramatically increased and it is now available for EFL learners as a free and promising material.

Many researchers have focused on the implementation of machine translation as a learning tool in EFL classes from different perspectives. While some are concerned with its impact on four skills, others focus on learner or teacher attitudes towards using a machine translation tool in EFL classes. Among those researchers who are concerned with its impact on writing ability, Chandra and Yuyun (2018) studied on how students made use of Google Translate in writing tasks and they found out that the students mostly regarded GT as an online dictionary and looked up words rather than translating full sentences or texts. In addition, Garcia & Pena (2011) search for its impact on beginners' writing skills, stating that most of the studies on MT are conducted among higher level participants. They conduct a two-phased experimental design. Firstly, the participants are required to write in their foreign language and then pre-edit it via a machine translation engine and on the next phase they write directly in their first language and then translate it into their foreign language. They find out that the participants communicate better via machine translation when they write directly in their foreign language. In other words they can produce more sentences when they get help from machine translation.

O'Neill (2011) follows a more complicated process in order to find out the effect of using machine translation in teaching French as a foreign language. The study consists of

three groups. The first group takes a training session on online translators and writes via an online translator; the second group does not take a training session but is permitted to use online translator; the third group is the control group and neither takes a training nor is allowed to use online translator. The products of the students are rated in terms of remaining grammar, spelling, comprehensibility, syntax, vocabulary and overall comprehensibility. The results suggest that the first two group did better compared to the control group in overall comprehensibility, content, spelling, and remaining grammar and the difference is statistically significant.

A study by Sangmin-Michelle Lee (2019) employs a different procedure in writing. The participants, who are intermediate or high intermediate, first write into their first language and translate it into their foreign language without any help from a machine translation engine. Later, they translate their writing using a machine translation engine and compare their own translation to the machine translation in order to decide on a final version of their text. The results show that using machine translator as a CALL tool reduces the level of lexico-grammatical errors and has a positive impact on students' revision ability. What is more, the students state in the interview that they are in favour of using machine translation in their writing classes.

A similar procedure is employed by Tsai (2019) and the results are more or less similar. The use of Google Translate in writing results in less grammatical and lexical errors providing students with more advanced level of vocabulary.

Another study by Groves and Mundt (2014) asks this question: Is machine translation a friend or a foe? The participants wrote an essay in their first language and these texts are translated into English without any post-editing. The translated texts are found to have errors but proficient enough to be admitted for a university. They conclude that although a machine translation engine cannot produce error-free texts, the imaginative ways of implementing

machine translation in Languages for Academic Purposes (LAP) classes will result in positive effects with the help of improving technology.

In addition to its impact on language skills, some researchers focus on the student's attitudes and perceptions for using machine translation. Most of the studies reveal that foreign language learners are in favour of using machine translation for their language classes. For example, in a study by Kumar A. (2012) every of the participants admits that they use machine translation in line with other studies by Alhaisoni & Alhaysony (2017); Bahri & Mahadi (2016); Jin & Deifell (2013); Sukkhwan (2014). A study by White and Heidrich (2013) reveals that the students are willing to use machine translation even though it is forbidden in their institute.

On the other hand, some instructors are of the opinion that using machine translation can be regarded as cheating, even as plagiarism. Correa (2014) argues that using an online translation in order to complete an assignment is a way of stealing language although the source idea is still the user's. Our students are supposed to write something on their own in their foreign language and they get marks for their ability to convey meaning in the foreign language. Thus, the researcher states, getting help from a machine translation engine is equal to cheating. She even regards looking up words on an online translation as dangerous in that it may result in phrase or sentence translation. According to her, a machine-translated text has no instructional value and should be graded with a bad mark. However, she suggests some activities can involve machine translation. The first one is to use machine translators for post-editing the written output. Here the students write a text directly in the target language and then post-edit it using a machine translator. The second idea suggests the students create four versions of a text. The first version is their writing in their first language. The second version is the traditionally translated writing, namely they translate their writing using dictionaries.

The third version is the machine-translated writing. Finally, they compare the second and the third version and edit them in order to create a final version.

Another suggestion from Correa is to refer to machine translation for pre-editing. The students can be required to pre-edit a text until they get an appropriate equivalence. In other words, the students write a text in their target language and translate it into their first language to see if they can convey the intended meaning. If not, they make changes in the first text until they can get the meaning. For example, the student writes 'There is a floor on the carpet.' And translate it to see if there is something wrong and pre-edit the text until he/she can get the accurate sentence.

Finally Correa suggests to teach students how to write more machine translation-friendly sentences. The differences between the languages can result in inaccurate translations. By creating different versions of a sentence, the students can be more aware of the metalinguistic structures of languages. In her article, she draws attention to the point that these activities with beginner level learners would "be behind the learners' zone of proximal development, or ZPD, and no learning can take place" (Correa, 2014, p. 15).

Among the researchers who are concerned with the use of machine translation as a learning tool are Ducar and Schocket. In their study in 2018, they state that "… L2 students consult the most widely used translation tool, Google Translate (GT), in spite of the fact that its use is frowned upon by second language (L2) instructors" (Ducar & Schocket, 2018, p. 779). They confront the instructors with an absolute reality: the machine translation technology will be gradually improved and it will challenge the instructors more and more. So the instructors are suggested to understand how this technology works and to find out suitable ways of teaching the students to make use of it efficiently.

Although machine translation has been proved to have such profits in learning a foreign language, there is something missing in such an automated technology: creativity (Saroukhil,

Ghalkhani & Hashemi, 2018). While human beings can convey meaning to cultural components of a simple sentence or can get a hidden gist of a single word, a machine cannot be creative enough to translate these unless it is programmed. Nevertheless, can machine translation help human being be more creative?

2.2 Creativity

Creativity has long been a subject for researchers from a number of perspectives. Since Guilford's presidential address in 1950 to the American Psychological Association about the need for research on creativity, it has gained popularity (Treffinger, Young, Selby and Shepardson, 2002). The fact that creativity is required in every field of our life ranging from solving a daily life problem to managing a company has enhanced the frame of these studies. As Stenberg and Lubart (1999) suggested that is because of her/his "creative vision of how to turn a company around" one is hired as a CEO, not because of her/his personality, knowledge or memory. S. Moran (1999) broadens the scale and claims that one can find creativity in

> … everyday cleverness, especially among children; the arts and sciences, with an abundant stream of paintings, dramas, theories, and concepts; business, with innovative products such as Federal Express's overnight delivery, 3M's Post-It Note, and Google; social interaction, most recently with Web sites like MySpace and Twitter; education as charter schools and non-school venues, such as children's museums, arise around the world; and public policy as countries try to govern and promote their cultural assets and intellectual capital in more systematic ways, such as England's cultural industries initiatives. (Moran, S., 1999, p. 74)

2.2.1. Definition. Creativity has been defined by many researchers in different ways and it can still be defined in an infinite number of ways as it has many faces in human behaviour influenced by several number of aspects, such as personality, environment, education and society. In other words, there is not a universal consensus on the definition of creativity. It has

such intertwined layers that there is still a controversy even over its definition. E. Paul Torrance (1988, p.43), who is one of the pioneers of creativity research, states that “creativity defies definition”. In addition, Kozbelt, Beghetto and Runco (2010) define creativity as such a rich subject that it is impossible for it to have only one static definition.

Nevertheless, some researchers have endeavour to define it by examining the creativity issue from a scientific point of view, the most well-known of whom are Guilford (1950, 1966), Torrance (1974, 1988, 1990), Amabile (1982, 1983, 1988), Stenberg & Lubart (1999) and Runco (1986, 2004). Kozbelt et al. (2010) categorised these theories of creativity in ten groups: developmental, psychometric, economic, stage & componential process, cognitive, problem solving & expertise-based, problem finding, evolutionary, typological and systems. As a result of this variety there is no consensus on the definition of creativity.

Maley and Bolitho (2015, p. 434) state that “creativity can be seen as the quality of being creative or as the ability to create.” And to create is defined by Krathwohl (2002, p. 215) as “putting elements together to form a novel, coherent whole or make an original product”. A group of educators, calling themselves C-Group, defines creativity as “Creativity is a quality which manifests itself in many different ways, and this is one of the reasons it has proved so difficult to define” (Maley & Peachey, 2015, p. 7). Nevertheless, the most frequently addressed expression defines creativity (Amabile, 1983; Feldman, Csikszentmihalyi & Gardner, 1994; Moran, 2010) as “ability to produce a novel yet appropriate solution to a problem or response to a situation”.

Although M. A. Boden (2001, p. 95) makes a definition of creativity in terms of being able to come up with a new, surprising yet intelligible and valuable ideas, she finds the term “new” ambiguous: “new for the person himself or new for the whole of human history”. Then she suggests to pick the option which meets the need of the research.

Teresa M. Amabile (1983) suggests that psychologists tried to define creativity under three main titles: creative process, creative person and creative product (p. 358). As stated by Amabile (1983), Wertheimer (1945), Koestler (1965), Newell, Shaw, and Simon (1962) saw creativity as a problem solving process in which the thinker became aware of the link between the problem and the possible solutions by means of being creative. Secondly, she states that some research tries to define the personality of creative people (Amabile, 2018). Guilford in his address to the American Psychological Association considered creativity in terms of creative person: "In its narrow sense, creativity refers to the abilities that are most characteristic of creative people" (as cited in Runco and Jaeger, 2012, p. 94). Amabile's (1983) last group of definition concerns creativity in terms of product, which is observable outcome or response (p. 358). Amabile, herself, thinks that a product-based definition is more straightforward and it is easier to observe and assess creativity in terms of product, hence she describes creativity as "the production of novel and useful ideas by an individual or small group of individuals working together" (Amabile, 1988, p. 126). The common aspects of a creative product in such product-based definitions are novelty, appropriateness and originality.

Kozbelt et al. (2010) focus on the necessity of moderation and pluralism in the understanding of human creativity. According to them, "to understand creativity in all of its richness, there is a need for moderation, where no one theoretical perspective is emphasized at the expense of others." (Kozbelt et al., 2010, p. 20). From this point of view, they classify the theories of creativity into ten groups: developmental, psychometric, economic, stage & componential process, cognitive, problem solving and expertise-based, problem finding, evolutionary, typological and systems.

As for the dictionary definitions, creativity in the Cambridge Online Dictionary (2019) is defined as "the ability to produce or use original and unusual ideas". Oxford Online

Learner's Dictionary (2019) says creativity is "the ability to use skill and imagination to produce something new or to produce art; the act of doing this". Lastly, Longman Dictionary Contemporary English Online (2019) explains the term as "the ability to use your imagination to produce new ideas, make things etc".

As it is clear, the common word in these definitions is novelty. As cited in Torrance (1988), Thurstone (1952) suggested that not the society but the thinker himself should render the product as novel to be counted as creative. Runco and Jaeger (2012) argue that if something is not new, or original in other words, it cannot be rendered as creative. Nevertheless, being original is not sufficient for an original idea or product, for it can be useless so in addition to being original, an idea or a product should be effective as well. Although some other labels are used instead of effectiveness, such as appropriateness, fit or usefulness, Runco and Jaeger state clearly that "original things must be effective in order to be creative" (Runco and Jaeger, 2012, p. 92).

2.2.2. Creativity and Joy Paul Guilford. J. P. Guilford (March 7, 1897 – November 26, 1987) was one of the leading figures in factor analysis in creativity research. Thanks to his presidential address to the American Psychological Association in 1950, research in creativity gained speed (Amabile, 1983; Barbot, Besançon and Lubart, 2011; Runco, 2004; Treffinger et al., 2002) although research in creativity had already started back in the first half of the century (Runco and Jaeger, 2012).

According to Guilford, creativity is a need for the spirit of the time. As cited in Pope (2005, p. 19) and Runco (2004) he says that "an unusually strong interest in the subject [Creativity] is an aspect of our Zeitgeist". Those times, the 50s, were the age of rapid changes in science, technology, society and military. Therefore, rapid and new responses to these changes challenged the humanity, which resulted in a dense inquiry in creativity. Especially the coming space age was a call for the "upsurge in interest in creativity" (Pope, 2005, p. 20).

Guilford (1950) defines creativity as a "combination of abilities" which can be found in every individual in different amounts (as cited in Rubinstein, 2003; Runco & Jaeger, 2012). According to him, every human being somehow can bear creativity in his/her acts. He argues that there are some reasons behind an individual's being creative or not in a specific period of time. He says that (as cited in Runco & Jaeger, 2012, p. 94) "Whether or not the individual who has the requisite abilities will actually produce results of a creative nature will depend upon his motivational and temperamental traits".

Guilford offers a theory on human intelligence named "The Structure-of-Intellect" (SOI; Guilford, 1956). He argues that human intelligence is a combination of many mental factors and it is not dominated by only one of them (Behr, 1970). In his Structure of Intellect Theory (SOI), Guilford (1968) makes a distinction between divergent thinking and convergent thinking (cited in Rubinstein, 2002; Kozbelt et al., 2010). Convergent thinking is the process leading to a convergent product, which is defined by Guilford as "Generation of information from given information, where the needed information is fully determined by the given information; a search for logical imperatives" (Guilford, 1970, p.158). On the other hand, divergent thinking is the process of creating a divergent product, which is again described as Guilford "Generation of information from given information, where the emphasis is upon variety and quantity of output from the same source; a search for logical alternatives" (Guilford, 1970, p.158).

He suggests that original and novel ideas or products are more likely to emerge when the creativity test allows divergent thinking. He is of the opinion that the more the participants are allowed to think farther from the starting point, the more they are likely to be creative (Kozbelt et al., 2010). Runco suggests that "divergent thinking ability is not equivalent to creative ability but it is indicative of the potential for creative performance. Hence, evaluating

the ideational patterns that are elicited by divergent thinking tests ostensibly helps researchers understand one component of the creative process" (Runco, 1986, p. 346).

Moreover, Guilford, as a factor analytic scientist, proposes that divergent thinking, which focuses on creative thinking, has four components: Fluency, flexibility, elaboration and originality (Hickey, 2001; Kim, 2006). According to Runco and Acar (2012), although divergent thinking had been mentioned in previous studies, it was Guilford who made the systematic connection between divergent thinking and creativity; therefore, divergent thinking tests now are investigating on fluency, originality, flexibility, and elaboration. For example, the most widely used creativity test Torrance Test of Creative Thinking (TTCT; Torrance, 1974), which is a pen-and-paper test, basically depends on the idea of divergent thinking and the responses of the test takers are considered in terms of fluency, flexibility, elaboration and originality (Hickey, 2001; Kim, 2006). In the verbal form of TTCT, the participants are required to write about before and after of a given picture (cause and effect relationship), how to improve a product or different and unusual ways of using a simple object. For instance, the participants list the unusual use of a cardboard box and their lists are scored in terms of fluency, flexibility, originality and elaboration. The quantity of the responses is the fluency score; the number of the categories in the list is considered as the flexibility score; the number of the infrequent responses in relation to the others' responses in the group is the originality score; the number of the details determine the elaboration score.

2.2.2.1. Fluency. In Guilford's words "Fluency is a matter of facility with which an individual retrieves information from his personal information in storage" (Guilford, 1966, p. 188). The only source for an individual to find out information in order to create something is his memory. Here Guilford makes a distinction between having the required information in mind, and being able to bring it back and use it. In other words, he suggests that there is a

distinction between “the memory abilities” and “divergent-production abilities” (Guilford, 1966, p. 188).

Three major types of fluency are listed by Guilford, which are "ideational fluency”, “associational fluency” and "expressional fluency” (Guilford, 1966, p. 188). The first one is creating ideas or thought as possible. The related tests ask the participants to make a list of ideas as rapidly as possible by just giving them two or three specific attributes, e.g. make a list of the animals that are soft and hairy. The associational fluency includes listing as many items as possible similar to an idea which is provided by the testers, e.g. list similar things to the thought of ‘hard’. The latter type, expressional fluency searches for the ability to produce a number of complex and organized sentences, e.g. make sentences that contains these three words: sea, cloud and girl.

Torrance (1990) describes fluency as the number of the ideas or thoughts that are listed by a single participant. The scoring demonstrates the ability of the individual of creating a flow of figural image. It may have high correlation with flexibility and in some research flexibility can be eliminated (Kim, 2006).

2.2.2.2. Flexibility. Flexibility is defined by Guilford as “Flexibility is a matter of fluidity of information or a lack of fixedness or rigidity. Novel output automatically implies new and unusual uses of retrieved information and also revisions of that information” (Guilford, 1966, p. 188). It is considered as the basis of originality.

Flexibility is divided into two main kinds. The first one is related to “class ideas: a facility to going from one class to another” (Guilford, 1966, p.188). A test in search of this kind of flexibility as the participant to think of unusual ways of how to use a thing, e.g. an invitation card and the participants need to go from one class to another in search of unusual answers, e.g. using it as a fan, clipping out some parts and use them for decoration, use it as a drip mat. The other type of flexibility is related to “shifts of meaning or reinterpretations in

the service of problem solving" (Guilford, 1966, p.188). The related tests ask the participants to produce titles for poems, stories or riddles; or to talk about the consequences of a particular event.

2.2.2.3. Originality. Originality is usually described in terms of novelty (Runco & Acar, 2012). The most infrequent ideas or thoughts are considered as the original ones. Originality of ideas are simply labelled by collecting ideas and identify the most infrequent ones. In order to be more reliable, having an inter-judge agreement is suggested by Runco and Acar (2012).

Ball & Torrance (1984) and Torrance (1990) describe originality in a similar way: the number of the unique ideas or thoughts that is stated by a participant, which shows the ability of an individual to produce unique and uncommon ideas. The scoring is based on normative data. Uniqueness is decided by compiling all the responses in the group and an idea or thought that is stated once in the group takes 1 and the others take 0 (Kim, 2006; Nusbaum and Silvia, 2011).

2.2.2.4. Elaboration. The last component of creativity is elaboration. It is regarded by Guilford as "a facility for adding a variety of details to information that has already been produced" and it is called as "finishing touches" (Guilford, 1966, p. 188). According to Torrance (1984, 1990) elaboration is the number of the additional ideas, which proves the ability of an individual to produce and elaborate ideas. The test of elaboration can be conducted by giving an outline of a plan for a fair and ask the participant to elaborate the plan in order to organize a successful fair.

Torrance (1998) offers the TTCT Norms of Technical Manual in order to help the researchers to get a Creativity Index (cited in Kim, 2009) and he suggests to get the average of the standardised scores of these four subtitles to produce an overall score for creative potential.

According to Guilford, such tests that assess the divergent thinking ability of the participants should ask them to produce something, rather than handing in a multiple-choice test. Secondly, reliability of such tests may be relatively low because of "the general instability of functioning of individuals in creative ways" (Guilford 1966, p. 201). Lastly, this kind of tests is not expected to have predictive validity as well because of the same reason.

Runco and Acar point at the above mention phenomena of creative measurement stating that an indicator of creativity is somehow a predictor and when a prediction is made there appears an uncertainty (Runco & Acar, 2012). When someone who has high level of creative performance in a divergent thinking test may have a lower ranking in the future. Similarly, a group or a nation may come up with creative thoughts in a divergent thinking test today but the same group or the same nation may slip up in another one in the future.

2.2.3. Creativity and language. As stated by Noam Chomsky in Language and Mind (1968) "A person who knows a language has mastered a system of rules that assigns sounds and meaning in a definite way for infinite class of possible sentences." (Chomsky, 1968). Namely, a language is a set of sounds and meaning with infinite combinations of letters, syntax and ideas in an infinite number of ways.

In terms of infinity, creativity has a significant role in language acquisition and learning for according to cognitive psychological researchers, creativity is "the ability to produce new valuable ideas and involves the use of various creative processes or thinking types such as exploratory, combinational and transformational thinking" (e.g. Boden, 2001, p. 96; Tin, 2015, p. 43). From that point of view, exploratory thinking, combinational thinking and transformational thinking have great importance to be counted as creative. Exploratory thinking is described as the ability of generating new ideas; combinational thinking of producing new valuable ideas in an exceptional way; transformational thinking of manipulating or transforming the existing ideas to create new ones. Thus, creativity, as a

combination of these three skills, must be on the top of the list in language acquisition as the most required ability.

In addition to the fact that our creative ability helps us to generate new utterances in any language, it also helps us to make meaning out of any utterances which are produced by others in any language. For instance, reading this following sentence “Daniel Boone decided to become a pioneer because he dreamed of pigeon-toed giraffes and crossed-eyed elephants dancing in pink skirts and green berets on the wind-swept plains of the Midwest.” (Fromkin, Rodman, Hyams, 2018, p. 10) you may not find any logic in it but you can still understand the meaning of it although it is the first time you encounter such a sentence.

2.2.4. Creativity and education. Torrance states, in search of a definition for creativity, that “When a person has no learned or practised solution to a problem, some degree of creativity is needed.” (Torrance, 1988, p. 57) in his book the Nature of Creativity. In other words, he suggests that a person should discover and discipline his/her creativity on his/her own in order to create a solution out of the provided opportunities. This idea raises a question: Can creativity be taught? Torrance gives an answer to this question with the help of his longitudinal research: “The truly creative is always that which cannot be taught. Yet, creativity cannot come from untaught.” (Torrance, 1988, p. 58).

Studies carried on the phenomenon of creativity have inevitably led the education researchers investigate about the creativity in classrooms. According to Beghetto (2010), psychologists agreed on the necessity of developing students’ creative competence as a curricular goal long time ago. In addition, Guilford (1967) and Torrance (1968) focused on the importance of raising the creative ability levels of school-aged children because they found out that creativity of many of the children in first and second grade had been suppressed by the school curricula.

Taking this claim into consideration, the activities in classroom can be arranged to flourish creativity. For instance, Gonçalves, Cabral, Campos and Schöning (2017) reported that creativity in classrooms can be sparked by referring to the senses of the students. They conducted a study with 100 students and searched for the effect of auditory and olfactory stimuli on the level of creativity in students' writing ability. It was found out that accompanied with the auditory cues, olfactory cues were significantly affective in the creative writing process.

In terms of technology impact on the level of creativity in classrooms, there are a lot of studies that prove technology to be a spark for creativity in education. Among them, Kolokoltsev, Iermakov and Potop (2017) found out that an educational application of technological tools in fine arts was more appealing in Physical Education than the conventional education tools.

In accordance with those findings, there is a number of studies on the phenomenon of creativity in EFL classes that suggest a positive correlation between the students' level of creativity and their class grades (e.g. Albert and Kormos, 2011; McDonough, Crawford and Mackey, 2015; Pishghadam and Mehr, 2011). Moreover, as Drapeau (2014) says "You cannot be creative unless you have something to be creative about." (p.147). In accordance with this idea, a study conducted by Wang (2018) states that creativity training infuses the students' fluency and originality in generating ideas.

In the 9th and 12th grades English curriculum which is published by Ministry of National Education (MoNE) in Turkey, students are viewed as "creative individuals who can produce language materials and tasks with the guidance of their teachers" (MoNE, 2018, p.10). As the expression defines, students can be creative once they are supported. Accordingly, Runco (2014) mentions that every one of us is born to be creative but we are not able to fulfil it. Aside from the effect of age, the environment also plays a significant role in

fostering creativity. According to Sir Ken Robinson (2001), cited in Lambirth, Goouch and Grainger (2005, p.10), one should be allowed to take risks and to have experiences with a "systemic strategy" in order to be creative. However, in a study which investigates the correlation between creativity, age, gender and academic achievements by Naderi, Abdullah, Aizan, Sharir, & Kumar (2009) it is revealed that creativity, age and gender are low predictors of academic achievement.

According to Amabile (1988), the features of the society can be determining in one's creativity. For example, evaluation, time pressure, previously-determined rewards, over-supervision, restriction of choices may impede creativity. Unfortunately, most schools are the places where one can find every one of those items in the list. In addition, the idea of risk-taking is one of the key issues of being creative; however, students are prevented from taking risks in the traditional education systems, thus from being creative. A student taking a multiple choice test cannot think creatively for he should think in the way that the curriculum requires him to do in order to be successful.

The traditional classroom environment with 75% of instruction hinders students from thinking creatively (Goodlad, 2004) and fosters convergent thinking. Moreover, as cited in Beghetto (2010, p. 451), Kennedy (2005) suggests that some teachers are observed to "habitually dismiss unexpected student ideas" for the fear of falling behind the plan, of chaos or as a personal choice. In a study conducted in Pakistan by Irfan (2018), the factors lying behind the lack of creative environment at schools are listed as the curriculum's lacking real problems, frequently used memorization technique, large-size classes, inadequate pre and in-service teacher trainings, and exam techniques.

Some of the researchers have put through some suggestions for teachers. For instance, Conti, Coon & Amabile (1996) suggests creativity is suppressed when the student is extrinsically motivated and it flourishes with intrinsically motivated students. Thus, comparison

and competitions among students or challenging tasks are demotivating activities that should be avoided in a classroom where creativity is a required component. Beghetto (2010, p. 457) mentions that creating self-efficacy of students helps them to generate "novel and adaptive ideas, solutions and behaviours".

Craft (2001) also suggests the teachers to apply the "Creative Cycle" method in their classes in order to foster creativity. The idea of creative cycle has first been put through by Wallas (1926) and the creative process is divided into four stages: preparation, incubation, inspiration or illumination and verification. Preparation is the stage of being disciplined and focusing on the subject. Incubation is when you let it go and do not think about the subject consciously. Inspiration is a spark that comes after the incubation stage. Lastly, one refines the outcome and the cycle starts again. According to Craft (2001, p. 19), this cycle suggests the students and the teachers:

> …to be open to possibility, the unknown and the unexpected; bridge differences – make connections between apparently unconnected ideas and integrate different ways of knowing (for example, physical, feeling, imagining); hold the paradox of form and freedom; hold the tension between safety and risk; be willing to give and receive criticism; be aware of the individual (Craft, 2001, p. 19).

2.2.5. Measurement and assessment. The fact that there is a controversy over the definition and the characteristics of creativity results in a contention in the measurement and assessment of it. The reliability and validity of the techniques have been a challenge for the researchers. Barbot et al. (2011) argue that:

> Due to the multiplicity of the conceptual approaches of creativity used at that time, the field of creativity assessment was viewed as experiencing a "mid-life crisis" with a problematic proliferation of assessment techniques showing lack of definition and limited educational applications. Most of these numerous techniques and new

> assessment tools have also been criticized for their weak psychometric properties or lack of up-to date norms to situate individual performance in developmental, gendered and cultural relevant groups of comparison. (Barbot, et al., 2011, p.59)

In other words, the vagueness in the definition of creativity results in the difficulty of creating new, reliable and valid techniques in creativity assessment.

Nevertheless, there are attempts to assess creativity. For example, there are some validated tests created by the pioneers, such as “the Alternate Uses” test by Guilford in the 1950s; “the Torrance Test of Creativity” (the TTCT) by Torrance based on the Guilford’s previous psychometric studies; “the Abbreviated Torrance Test for Adults” (ATTA) by Goff and Torrance (2002); “the Abedi Test of Creativity” (ATC; Abedi, 2000) and many other divergent thinking tests.

Generating a creativity test proposes another challenge for the researchers: scoring. There are some different ways of scoring in such divergent thinking tests which are mainly concerned about the fluency, flexibility, originality and/or elaboration. The most common one is generated by Wallach and Kogan in 1965 as cited in Silvia et al. (2008). The test requires the participants to write about unusual uses of an objects, e.g. a brick, a cardboard, a knife and the raters count the responses of each participants in order to determine the score for fluency and then they search for the responses that are stated just once in the study group in order to determine the score for uniqueness, in other words originality. The statement that is mentioned once gets 1 point and the others a 0. Therefore, the results indicate two results, one for quantity, and the other for quality of divergent thinking (Silvia et al., 2008).

However, counting every response in a study with large number of participants requires a large amount of effort and time (Nusbaun & Silvia, 2011; Silvia et al., 2008). The problems in this scoring method are listed by Nusbaum & Silvia (2011) in two main items.

The first problem is that the uniqueness scores are depended strongly on the size of the sample. In other words, as the size of the sample increases, the mean score of uniqueness decreases. While a response is labelled as creative in a smaller group, the same response may be labelled as not creative in a larger group. Therefore, it is getting harder to be detected as creative in larger-sized groups.

The second problem is about the correlation of fluency and uniqueness scores. As the sample is getting larger the correlation is increasing. Nusbaun and Silvia (2011) argues that the correlation in a variable reanalysis of Wallach and Kogan's data by Silvia in 2008 is calculated as r = .89 and they question whether this highly correlated two variables are actually two district variables.

In addition to this approach which tries to count every responses in terms of fluency and uniqueness, Amabile suggests a product-based approach, the name of which is the Consensual Assessment Technique (the CAT, Amabile, 1982). The CAT is defined by Hennessey and Amabile (1999, p. 347) as "a technique used for the assessment of creativity and other aspects of products, relying on the independent subjective judgements of individuals familiar with the domain in which the products were made". For further information see section 2.2.5.2.

2.2.5.1. Reliability and validity of creativity tests. Reliability is described by Dörnyei (2011, p. 50) as "… reliability indicates the extent to which our measurement instruments and procedures produce consistent results in a given population in different circumstances". As Runco and Acar (2012) suggests that especially the interrater reliability is of vital importance if there is a subjective scoring system, which is the case in most of the creativity research. As cited in Runco and Acar (2012), the interrater reliability coefficient is found to be above .90 in Meeker's research on Structure of the Intellect Learning Abilities Tests and it is reported in Urban and Jellen (1996) work as .90 and Wallach and Kogan (1965) .92. It is clear that the

interrater reliability is straightforward but the researcher is suggested to decide on well about how much information should be shared with the raters and to be consistent in choosing the best judges.

Hocevar and Michael (1979, p. 917) offer a percentage formula in order to calculate validity, and their study gives information about coefficient alpha reliability:

> … the discriminant validity of subjectively determined originality scores is enhanced when the scores are divided by the total number of responses (yielding a percentage score). However, when the percentage formula was used to derive flexibility scores from the Torrance Tests or an originality score based on statistical rarity from the Alternate Uses test, the coefficient alpha reliabilities dropped to near zero. This finding suggests that the utility of the percentage formula is limited to subjectively determined originality scores (Hovevar & Michael, 1979, p. 917).

In short, when the fluency scores are statistically controlled, the originality results gets more unreliable. However, in their study with gifted and non-gifted children, Runco and Albert (1985) oppose this idea and offer that this is not consistent with every tests or every sample.

As for validity, Runco and Acar ask the related question: "Are you testing what you think you are testing?'' and put validity up against bias (Runco and Acar, 2012, p. 67). The researcher should be aware of the fact that the test can be influenced by something which is not planned or foreseen. For example, in creativity research some participants may have low verbal skills, which may result in low creativity scores even if they have creative abilities.

Some other potential biases can be listed under the title of discriminant validity. The creativity researchers have long been trying to find if there is a correlation between the intelligence and creativity, relatively convergent thinking and divergent thinking. While some thinks that there is a relation, some other oppose. Torrance (1966) suggests that "some level

of general intelligence is necessary but not sufficient for creative work" (cited in Runco and Acar, 2012, p. 67). However, as creativity is an intertwined subject, it is being tested with different groups with different level of intelligence in order to define if creativity tests really test the creative thinking or the effect of intelligence on problem solving.

As for predictive validity, it is defined as the degree of relation among the tests and the criteria. There are several endeavours in the creativity research field to get stronger relations between the criteria and the tests. For example, Hocevar (1981) states that among ten different categories of criteria of creativity – "tests of divergent thinking, attitude and interest inventories, personality inventories, biographical inventories, teacher nominations, peer nominations, supervisor ratings, judgments of products, eminence and self-reported creative activities and achievements - the self-reported creative activities and accomplishments are the most defensible technique for selecting creative individuals" (p. 450).

In response to the criticisms on the validity of creativity tests, Torrance, who is the creator of the TCTT, suggests that the creativity tests are the most investigated tests among the educational tests. Hence, the administration and the evaluation processes have been experimented for many times to provide precious data about being reliable and valid (Torrance, 1988).

2.2.5.2. The Consensual Assessment Technique. "The Consensual Assessment Technique (the CAT; Amabile, 1982) is based on the idea that the best measure of the creativity of a work of art, a theory, a research proposal, or any other artifact is the combined assessment of experts in that field. Unlike other measures of creativity, such as divergent-thinking tests, the Consensual Assessment Technique is not based on any particular theory of creativity" write Baer and McKool (2009, p. 1). Unlike the previous test techniques which require an evaluation by just the experts in creativity, the CAT lets the evaluators, who are experts in a similar domain to the products, judge subjectively either relying on a set of

criteria or on their own sense of creativity (Baer, Kaufman & Gentile, 2004). In the CAT, the products of a person or a group of people are compared and ranked order from not-at-all creative to high creative (Hickey, 2001; Barbot, Besançon and Lubart, 2011; Silvia, Martin & Nusbaum, 2009). The underlying scope of this technique is the fact that there exist no absolute norms while assessing the creativity of production, thus the products should be ranked order relatively compared to one another (Hickey, 2001; Said-Metwaly, Van den Noortgate, Kyndt, 2017).

The CAT does not assess or label a person's creativity, instead it just assesses the product. Therefore, the studies which imply this technique simply want the participants to produce something, e.g. a piece of writing, by giving a prompt and then some experts assess the product independently (Baer et al, 2004). For instance, in a study conducted by Baer (1994), the participants are given a simple line drawing of a boy and a girl and are asked to write a story in which these boy and girl take part. Then the experts rank the stories relatively from 1 to 5 in terms of their own sense of creativity.

According to Silvia et al. (2009), creativity assessment needs a lot of time, effort and personnel. Thus, he suggests to implement such holistic techniques as the CAT in order to get rid of these challenges which hinder the researchers from conducting research in the field of creativity.

Baer et al. (2004) conduct a study by extending the CAT to assessing nonparallel products which are written under uncontrolled conditions in terms of creativity. Their idea is to implement the CAT with the help of 13 expert judges, who are writers or editors, teachers and psychologists. The experts are informed in the very beginning of the study that the only criteria in rating is creativity and as they are experts in their own field there is no need for them to express and defend their creativity definition. They are free to make their own definition of creativity.

Then the experts work on the products independently and firstly they divide 103 stories, 103 personal narratives, and 102 poems - every of which was written by 8th grade students not actually for this study and presented to the researchers by the students themselves as their most creative products - into three groups: low, medium and high creative. Then they compile each group in six levels of creativity.

The results show that there is a high level of inter-rater reliability: "0.940 for the stories, 0.957 for the personal narratives, and 0.868 for the poems" (Baer et al., 2004, p. 155). They argue that this high level of inter-rater reliability is due to the large number of creative products and the experts. In studies that implement the CAT in 1983 and later secondary analysis in 1996 by Amabile, who include such large numbers of experts, a similar degree of inter-rater reliability is calculated; however, as the number of the raters and the products decreases in many other research, the degree of agreement between the raters decreases relatively (as cited in Baer et al, 2004).

After obtaining such a high degree of inter-rater reliability, they conclude that the CAT can give reliable and valid results with products written even under uncontrolled conditions, which can open up new dimensions for creativity research.

As stated by Barbot et al. (2011), after inter-rater reliability is determined to be satisfactory, the second step is to take the average of the scores that are given by the judges to each product to derive an overall score for creativity.

Henessey defines the rationale behind the CAT as "a product or response is creative to the extent that appropriate observers agree that it is creative" (Henessey, 1994, p. 193) and he argues that there is a great number of data which proves that the CAT is reliable and has construct validity. On a study that she implements the CAT and searches for the mechanisms that underlie the responses of judges, she comes up with three findings and reports as "Findings were that: (a) judges were able to reliably assess not only the creativity of a

finished product but also the creativity of the process that went into producing that product; (b) ratings of process and product creativity (as well as a variety of other dimensions) tend to be highly correlated; and (c) information about the age of a creator can also significantly affect judges' subjective assessments" (Hennessey, 2009, p. 193).

2.2.4.2.1. Reliability and validity of the CAT. Baer and McKool (2009) state that in order to get reliable results in the CAT, having a group of independently working judges is needed. Then the inter-rater reliability is simply measured by "Cronbach's coefficient alpha, the Spearman-Brown prediction formula, or the intra-class correlation method" (Baer & McKool, 2009, p. 5). In the field of creativity research, correlation coefficients between .70 and .80 are believed to indicate a strong degree of agreement among raters (Conti et al, 1996; Baer et al., 2004; Baer and McKool, 2009).

Amabile (1982) mentions that if a product or a response is agreed to be creative by independent observers, it means that it is creative, which actually means that the inter-rater reliability is also an evidence for the construct validity (Hickey, 2001).

As cited in Baer and McKool (2009), Amabile has conducted a number of validation studies and she finds out that the CAT has discriminant validity. For example, in one of her studies in 1983, she asks the raters to judge creativity accompanied with some other attributes of the products in question, such as technological goodness (correlation with creativity = .13), expression (correlation with creativity = - .05), neatness (correlation with creativity = -.26) and she finds out that the judges rates the products in term of creativity much or less in an agreement while they contradict in other attributes.

The idea that creativity is domain specific and it is not a general trait is tested via the CAT as well. The correlation calculations among the scores of subjects' products in different domains show nearly no traces of positive or negative correlations (Baer and McKool, 2009). As it can be stated in other words that the CAT assesses just the product in a specific domain

in relation to the others in a specific group or to his/her previous products and the fact that a product is rated as highly creative in a study among the others cannot be generalized to other domains. In short, the fact that a person can create a highly creative story does not mean he/she can be a creative chef.

In order to get rid of some potential biases, such as social environment and the effect of the controlled conditions of a study, in research with the CAT, Conti et al. (1996) suggests not to create a contestant environment. Instead, fostering intrinsic motivation primarily via raising interest, satisfaction and joy can result in more creative works. According to her, people are more creative when they are intrinsically motivated.

2.2.4.2.2. Inter-rater reliability in the CAT. The Consensual Assessment Technique requires raters to rate a number of products independent from each other. Thus, inter-rater reliability is an important feature for research which implements this technique. According to Conti et al. (1996), correlation coefficients between .70 and .80 are indication of a strong agreement among the raters in creativity research. As cited in Baer et al. (2004), a flow of research by Amabile (1983, 1996), Baer (1993, 1998), Hennessey and Amabile (1999) and Runco (1989) have satisfactory inter-rater reliability scores, ranging from .70 to even more than .90.

2.3. Surface Structure Taxonomy

Errors are inevitable parts of learning a new language and they have attracted attention of a number of researchers who want to yield precious insight into language accusation. Among them are Dulay, Burt and Krashen have a remarkable place with their book Language Two (1982). They put a light on what is behind learning a new language ranging from the effect of age to motivation or analysis of errors.

According to the authors of the book (Dulay, et al, 1982, p. 5)

> … the most surprising finding in L2 acquisition research concerns the errors second language learners make. For several decades, linguists and teachers assumed that most second language learners' errors resulted from differences between the first and second languages. This was the basis of the long-popular contrastive analysis theory. Now, researchers have learned that the first language has a far smaller effect on second language syntax than previously thought. Studies show, for example, that only 5% of the grammatical errors children make and at most 20% of the ones adults make can be traced to crossover from the first language.

In order to classify errors, the Surface Structure Taxonomy was introduced by Dulay, Burt and Krashen in Language Two (p. 150) to help the researchers understand the logic behind the learners' errors. Since then it has long been consulted in a number of research which are about grammatical accuracy (Btoosh, 2011; Costa, Ling, Luís, Correia & Coheur, 2015; Garrido & Rosado Romero, 2012; Lee, 2019; Maniam & Rajagopal, 2016; UniKl, Saad & Sawalmeh, 2014).

Surface Structure Taxonomy divides the errors in four main groups: omission, addition, misformation and misordering.

2.3.1. Omission. Omission errors are identified as "the absence of an item that must appear in a well-formed utterance" (Dulay at al., 1982, p. 154). Most grammatical morphemes, such as noun and verb inflections (plural –s, apostrophe –s, -ed, -ing); articles (a, an, the); verb auxiliaries (am, is, are, will); prepositions, are more likely to be omitted than the content morphemes. Dulay et al. (1982) state that the omission of content words are most probably because of the lack of vocabulary knowledge.

2.3.2. Addition. Addition, as the name suggests, is "the presence of an item that must not appear in a well-formed utterance" (Dulay at al., 1982, p. 156). Three types of addition

errors are recognized: double marking, regularization, simple addition. Double marking results from failing to delete unnecessary items, such as the errors in those sentences:

I didn't went to school yesterday.

He doesn't knows my name.

You don't hardly eat.

In the first sentence, the main verb is uncessarily in the past form and in the later example, the –s is erronously added. In the last sentence 'don't' should be omitted.

The second type of addition is regularization. It consists of erronous additions to an exceptional item, such as fish-fishes, ate-ated, children-childerns, a book – a this book. The simple addition type is an umbrella term. Those additions which do not fit into the previously stated types are in this group.

2.3.3. Misformation. Misformation errors are defined as "the wrong form of the morpheme or structure" (Dulay et al., 1982, p. 158). Unlike omission errors, the learner suplies something which is inaccurate, as in the example: he eated dinner. Misformation errors are divided into three subgroups: regularization, archi-forms and alternating forms. In regularization group a regular marker is used instead of an irregular marker: gooses – geese, hisself – himself, childs – children. Archi-form errors include those in that a learner picks a form among a class of froms and over-use it as a representative. For example, the learner always use 'that' instead of picking the accurate one among 'this, those, these' or one may use only infinitive form of verbs consistantly instead of using gerund although needed: she finished to watch TV. The last group on misformation is alternating forms, which is the misuse of a member of a class of forms: he for she, this for these, her for she, saw for seen.

2.3.4. Misordering. Misordering errors are defined as "the incorrect placement of a morpheme or group of morphemes in an utterance" (Dulay et al., 1982, p. 162), such as 'I study all the time lesson'. Some of the errors in this category stem from a direct translation

from native language (Dulay et al., 1982). It is probable that a Turkish student writes ‘I every day school go.’

Chapter 3

Methodology

3.1. Introduction

The methodology chapter is a dense chapter consisting nine main items: aim of the study and research questions, research design, context, participants, data collection instruments, data collection process week by week and the data analysis procedure.

3.2. Aim of the Study and Research Questions

It is clearly seen in the literature review part that the research on MT is numerous. However, its implementation in the foreign language education field has been subject of few research with especially beginner level of participants. What is more, the effect of machine translation as a learning tool on creativity has rarely been a research subject. Accordingly, the aim of this research is to put light on the potential of MT as a MALL tool in foreign language classrooms in terms of creativity while searching for answers for the following research questions:

1. How does implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, elaboration and originality?

2. How does implementing machine translation in writing activities in English affect the participants' grammatical and lexical knowledge?

3. What are the perceptions of high school EFL students on the use of machine translation in writing activities in English classes?

3.3. Research Design

As mentioned by Dörnyei (2011, p. 42) the studies that combine quantitative and qualitative methods are named as mixed method studies. In their book "Handbook of Mixed Methods in Social and Behavioral Sciences", Tashakkori and Teddlie (2010, pp. 8-9) state that "methodological eclecticism" is one of the characteristics of mixed method research,

which allows the researchers to combine qualitative, quantitative and mixed methods in a synergy in order to investigate a subject more deeply.

Accordingly, Treffinger, Young, Selby and Shepardson (2002) state that:

> The complex and multidimensional nature of creativity cannot be captured effectively and comprehensively by any single instrument or analytical procedure. Systematic efforts to understand creativity require a well-planned process of studying individuals or groups, including both qualitative and quantitative data. (Treffinger et al., 2002, p. xi)

Therefore, this study bears an exploratory mixed method design, which is described by Caruth (2013) as a type of mixed method research in which quantitative data is gathered first and then the qualitative data in order to enhance the quantitative findings.

In addition to these, Morse (2009) mentions that there are "two main components of a mixed method research: the core and the supplemental" (p. 23). The core component of this research is gathered through quantitative method. The texts written by the participants as pre and post-tests are evaluated in terms of overall creativity by implementing the Consensual Assessment Technique (the CAT; Amabile, 1982) by five randomly chosen English teachers and in terms of fluency, flexibility, originality and elaboration by the researcher herself. Secondly, in order to investigate the second research question, the grammar and the lexis errors of each student both in their pre and post-test products are analysed to get another quantitative data set.

The conjunction point of the quantitative and the qualitative data is the semi-structured interviews with the volunteer participants after the implementation in order to find out the perceptions of them on the use of MT as a learning tool, which is the issue of the third research question. According to Scott and Morrison (2007, p.134), "semi-structured interviews let the researchers to deeply probe into the subject by means of being more flexible about the questions".

The validation in mixed method design research is mainly about the conjunction point of the qualitative and the quantitative data. According to Caruth (2013), the researchers should "provide an overt discussion and assessment of how the findings have been integrated from both quantitative and qualitative designs and the quality of the integration. This discussion will provide readers with an understanding of whether the inferences are in harmony with the research objectives and whether they make a contribution to the body of knowledge" (Caruth, 2013, p. 116). In order to provide the aforementioned overt discussion and assessment, five questions have been pre-determined for the semi-structured interviews. The reason behind choosing these questions is to support the findings of quantitative data and to put a light on the students' perceptions about using Google Translate. The questions are as follows:

1. What do you think the advantages and the disadvantages of using machine translation in the writing tasks were?
2. Do you think using machine translation has an effect on your creativity?
3. Do you feel like you have learnt new things while using machine translation?
4. Did you face any technical problems while you were using machine translation?
5. Had you used machine translation before we started that study?

Apart from being a mixed method research, the study is designed as a quasi-experimental research. The rational grounded behind is that it is conducted in a high school and participant assignment cannot be practiced randomly. Quasi-experimental design is mostly applied when it is not practical to assign participants to the groups in educational research although there are some controversy on the validity issue when compared to the true experimental design (Gribbons & Herman, 1996). Since the study is a quasi-experimental research, one of the classes among the present seven 9^{th} grades is chosen randomly. The class

consists of 35 students and both their parents and they have given written consent to participate in the research.

It is noteworthy that the study is a one-group, pre and post-test research design. As a quasi-experimental research, only a class of 9th grades is assigned as the study group. The study has not an experimental group. According to research designers, one group pre and post-test design has some threats to the validity of the study. Campbell and Stanley (1963, p. 7) state that the design can "jeopardise internal validity". They make a list of sources which affect validity: history, maturation, testing and instrumentation. History concerns the events occurring during the treatment. That can be anything that may affect the variables, such as a war, season or institutional schedule. Maturation is any physical or physiological variation during the time between the pre and the post-test. The participants might get old, hungry, angry, bored or might fall in love with the researcher. The next source is testing, which concerns the fact that the participants do better in post-test inevitably because they retake a same or similar test. The last one is the instrumentation, which is described by Campbell and Stanley (1963, p. 9) as the "instrument decay". It refers to the any changes in the research instrument. For instance, the observers might get more skilful after the pre-test or the interviewer might get used to the participants until the post-test.

In an endeavour to get rid of these possible biases as much as possible, the treatment period is kept short: seven weeks. The fact that there is only one group and the teacher is the researcher as well minimises the risk of extraneous effect on the variables. As for the instrumentation item, the products in the pre and the post tests are ranked by five English teachers in term of the CAT. The number of the participants is just 35, which is expected not to cause any fatigue in the raters.

3.4. Context

The present study took place in a state Anatolian high school in Bursa, Turkey in the first term of the 2019-2020 educational year with the 9th graders. In addition to four hours of compulsory English, two hours of elective English course was available for the 9th graders in the school. The study was conducted in the elective course for nine weeks in line with the curriculum which had been announced by the Turkish Ministry of Natural Education [MoNE] in 2018. The focus in the curriculum is on the communicative competence accompanied with function and four language skills (MoNE, 2018). Mobile assisted learning (MALL) is also praised in the curriculum and the teachers are suggested to integrate technology in their classrooms (MoNE, 2018).

The English classes in this school were designed in accordance with the 9th and 12th Grade Curriculum by MoNE (2018), which refers to the Common European Framework of Reference for Languages (CEFR). The 9th graders were regarded as A1 (Basic Users) and the objectives of writing activities were determined accordingly. The objectives for the first three months during when the present study was conducted were as follows:

- Students will be able to introduce themselves and their family members.
- Students will be able to ask and answer about their personal belongings.
- Students will be able to write simple sentences and phrases (a postcard, an e-mail or a hotel registration form).
- Students will be able to fill in a chart comparing cities in different countries/Turkey.
- Students will be able to describe different environments in simple sentences and phrases.
- Students will be able to write their opinions on a blog.

- Students will be able to write short text messages to invite their friends for a movie (pp.23-25).

These objectives were taken into account in designing the pre-post tests and the process in between for the students participating in the study were intended to keep up with the other 9th graders in the school.

3.5. Participants

The present study was conducted in a state high school by the researcher herself. Thus, the participants were high school students and the researcher was the instructor of the participants.

3.5.1. The students. As a part of conducting a quasi-experimental research, one out of seven 9th grade classes in the school was chosen randomly and the study was held with the help of the members of that class during the autumn semester in 2019-2020 educational year. The selected group consisted of 35 students (21 females, 14 males), whose parents had given a consent. Six of them (3 females, 3 males) volunteered in the semi-structured interview.

The participants' background information was collected with a form (see Appendix 1). The 31 of the participants were fourteen and 4 of the participants were fifteen years old. They were all born in Turkey and had not been abroad except for short-term family visits. They are native speakers of Turkish and two of them are bilingual, Turkish and Kurdish.

They had been studying English as a foreign language for five years at school before they started high school. They took the placement test on the DYNED (Dynamic Education) application via their mobile phones at home, which is provided by the Ministry of National Education for every student in Turkey and they were all labelled as beginners. Most of them stated that they usually have difficulty in learning English because they did not have the opportunity to use it affectively in their daily life. A few of them stated that they felt

motivated to learn English. The rest did not include English in their future plans, thus they were not motivated.

3.5.2. Role of the researcher. The researcher was also the teacher of the class. She graduated from English Language Teaching department in 2009 and had been teaching English for ten years in different state schools in Turkey. She had taught 9th graders for three years in that high school and was familiar with both the curriculum and the attitudes of 9th graders as foreign language learners. However, the study was conducted in the autumn semester and the students were newcomers and the researcher first met them then just before starting the research having no prior knowledge or prejudgement against the participants.

3.5.3. The raters. The study conducted the CAT in order to search for the overall creativity in the participants' written products. The technique requires a consensus among independent raters who are experts in the same domain as the products. Therefore, five randomly chosen English teachers (2 females, 3 males) rated the products both after the pre-test and the post-test. All of them have at least a ten-year of experience in teaching English to beginner level of students in high schools.

3.6. Data Collection Instruments

As it is a mixed method study, it requires to have both quantitative and qualitative data collection instruments. The quantitative data set includes the products which students wrote in pre and post-tests. The qualitative data is represented by the recordings of the semi-structured interviews with 6 voluntary students.

3.6.1. Google Translate. Implementing MT in writing tasks in EFL classes forms the core of the present research so one of the most frequently used mobile applications, Google Translate (GT, https://translate.google.com.tr/), was decided on to be the instrument of the study. In addition, GT is a free and accessible application for every operating system used for

mobile devices and has an easy interface. It can also offer instant translation, which helped with the time management during the implementation.

GT provides dual translation in more than a hundred languages. However, the methods it uses while translating and how accurate the translation are out of the scope of this research. The only thing that is concerned is the students' perceptions on MT, whether using GT affects the creativity of the participants and the impact of it on their lexical and grammatical knowledge.

Thus the students brought their mobile phones to the classroom and operated GT application during the treatment process. Due to the technological deficiencies in the classroom, it was made sure that every student had a personal Internet access. The researcher shared the Internet with those who needed.

3.6.2. Semi-structured interview. The interview was held one by one with the voluntary students (N= 6). The interviews were recorded and interpreted by the researcher. The interviewees were asked at least five questions, which are as follows:

1. What do you think the advantages and the disadvantages of using machine translation in the writing tasks were?

2. Do you think using machine translation has an effect on your creativity?

3. Do you feel like you have learnt new things while using machine translation?

4. Did you face any technical problems while you were using machine translation?

5. Had you used machine translation before we started that study?

The reason to ask these questions was to probe into the findings in the quantitative data analysis and answer the first question of the study.

3.7. Data Collection Process

The participants were ninth grade students and it had been four weeks since they started high school when they took the pre-test. Thus, they had a limited number of English

courses in their new school with their new teacher, who was the researcher as well. Taking the pre-determined writing objectives for this course into consideration, they were supposed to know how to introduce themselves and their family members how to ask and answer about their personal belongings and how to write simple sentences and phrases (a postcard, an e-mail or a hotel registration form) by the time the study had started. The data collection process is tabled in the Table 1 in detail week by week.

Table 1

The data collection process

Product	Time	Aim
Pre-test – a text	30 minutes	The students wrote a text about a stick-man without any help from Google Translate.
Training session	40 + 40 min.	The students learned about how to type machine-translation-friendly sentences.
Week 1 – a paragraph	40 + 40 min.	Unusual Uses - The students described their dream room including examples of unusual uses of a clock (fluency, flexibility and originality).
Week 2 – a list of questions	40 + 40 min.	Asking Questions - The students created as many questions as possible without giving answers to them (fluency, flexibility and originality).
Week 3 – a four-paragraph text	40 + 40 min.	Revision – The aim was to make a revision of the previous two units (elaboration).
Week 4 – a paragraph	40 + 40 min.	Product improvement – The students wrote about one of their favourite belongings and added some features

		in order to improve it (fluency, flexibility and originality).
Week 5 – a dialogue	40 + 40 min.	The students tried to be fluent and original while making excuses for a friend who invited them to cinema (fluency, flexibility and originality).
Week 6 – a table	40 + 40 min.	Just suppose – The students supposed that they were a director and wrote about their own movies (fluency, flexibility, originality and elaboration)
Week 7 – a two-paragraph text	40 + 40 min.	Expressing opinions – The students expressed their opinion about their friends' movies (fluency, flexibility, originality and elaboration).
Post-test	40 min	The students wrote a text about a blurry image of a human first without getting help from Google Translate for 20 minutes. Then they edited their writings with the help of GT for the next 20 minutes.
Interview	10 min (for each interviewer)	6 volunteers were interviewed about the use of Google Translate as a learning tool in writing activities.

3.7.1. The implementation of the pre-test and the post-test. In most of the studies on machine translation, the participants write or translate only one text using machine translation and these written products are evaluated in terms of grammar, lexis, strategies or the perception of the participants is inquired (Alhaisoni and Alhaysony, 2017; Bahri and Mahadi, 2016; Calis and Dikilitas, 2012; Case, 2015; Clifford, Merschel and Munné, 2013; Garcia and Pena, 2011; Groves and Mundt, 2014; Jin and Deifell, 2013; Jolley & Maimone,

2015; Kliffer, 2005; Kumar, 2012; Nino, 2008; Nino, 2009; Tsai, 2019; White and Heidrich, 2013).

One of them was conducted by White and Heidrich (2013) and they required the participants to write in their first language, English in that case, and then translate it into their foreign language, which was German, by using machine translation and lastly they were asked to make necessary editing to create a final work. Their aim was to see the participants' strategies in translation tasks.

In another study by Garcia & Pena (2011), the participants were divided into two in the writing process. One of the groups wrote by means of machine translation assistance and the other group wrote directly in their second language in order to see which way was more productive in written communication. They found that participants could produce more number of words by using machine translation and also the quality of the writings were higher with MT assistance.

In an experimental study, which is very unique, on the effect of machine translation on students' writing ability by O'Neill (2012) the participants were allowed to use machine translation in the instruction process while it was not allowed in pre and post-test administration in order to find out their writing ability without an access to a machine translator.

In the present research, a combination of these techniques was implemented. The technique by O'Neill, which was a pre post-test designed research like the present one, was administered in the pre-test process and the participants did not get any assistance from a class mate, a dictionary or machine translation in order to find out their sheer writing ability and creativity without an assistant tool (see Appendix 2).

As for the post-test, the students were asked to write a text directly in English in twenty minutes by looking at a picture of a blurry image of a human which is very similar to

the stick-man in the pre-test (see Appendix 3). Then they consulted to Google Translate in order to pre-edit their writing and create a final version in another 20 minutes. These final versions were regarded as post-test products (see Appendix 4).

As for creativity aspect of the pre and post-test, according to a study conducted by Amabile (1988), 74% of the participants mention that freedom is among the qualities of an environment which foster creativity. In other words, these people are of the opinion that they can be more creative if they are free to have control over their own work, making their own decision or how to manage their study. Thus, in the pre-test the participants were provided just with a drawing of a stickman, which was an adaption of a study by Baer (1994), and a blurry figure of a human-being in the post-test as a prompt to start writing and they were told that they were free about whatever or however they wanted to write about these pictures. On the other hand, in the same study by Amabile (1988), 52% of the participants state that they need sufficient sources to be creative. Namely, they feel themselves more creative when the necessary resources such as information or equipment are accessible. Therefore, the students were just required to imagine and to write about these pictures on a piece of paper in the allotted time and they were not allowed to add any details by drawing in case they might feel inhibited just because they did not have necessary equipment or drawing skills.

3.7.2. The instruction process. After the pre-test was administered, the participations got a seven-week instruction in their elective English lessons. They took two forty-minute lessons every week on how to write and be creative in their foreign language as beginner learners with the help of Google Translate. As the study was conducted in a state high school, the objectives were already determined by the Ministry of National Education but the activities were designed by the researcher herself by taking these objectives into account.

In the present study, Google Translate was utilised as a MALL pre-editing tool. The types of pre-editing tasks with machine translation are categorised by Shei (2002) in three

groups. The first way is to write in the first language (L1) and get a second language (L2) output. The objective is to modify the L1 text until you get a perfect L2 text. The second way is that a native speaker of the target language (L2) writes a text and machine-translates it into L1. The objective is to modify the L2 text until you get a perfect L1 text. The third way is quite similar to the second one but this time a learner writes a text in L2 and gets an output in L1 so that he/she modifies the L2 text until the L1 translation is perfect. The way the participants followed during the instruction process and the post-test is the third one in this study. First, they wrote in their L2 (English in this study) on their own and then they translated it into L1 via Google Translate to see if they could transmit the intended meaning. If so, they put a tick. If not, they pre-edited their English until they got an accurate Turkish translation. While pre-editing, they again got help from Google Translate.

It is noteworthy that this process was not a creativity training programme. The primary aim of the process was to help students to get familiar with the use of Google Translate in their regular writing activities. Nevertheless, the subjects of some of the activities were designed in terms of divergent thinking skills, which are evident in the verbal form of Torrance Test of Creative Thinking (Torrance, 1974), in order to help the students to flourish their creativity with the help of machine translation. Those divergent thinking skills are listed by Cramond, Matthews-Morgan, Bandalos & Zuo (2005) as "asking questions, guessing causes, guessing consequences, product improvement, unusual uses, and just suppose" (p. 284).

3.7.2.1. Training session. Before starting writing with Google Translate the participants were given a one-hour training session on how to use Google Translate and how to type machine translation-friendly sentences (see Appendix 5). The participants first discussed the advantages and disadvantages of machine translate and then machine-translated some example sentences both from English to Turkish and from Turkish to English. The

example sentences were determined beforehand in order to serve a practical objective of the session. For instance, they translated a poem and an idiom to find out that machine translation is not good at getting the cultural issues. They also realized that some of the words did not necessarily have a strict translation, such as yellow-blonde, kara-siyah. The role of the punctuation and capitalisation was emphasized with some striking examples (Retrieved from https://translate.google.com on November 23, 2019):

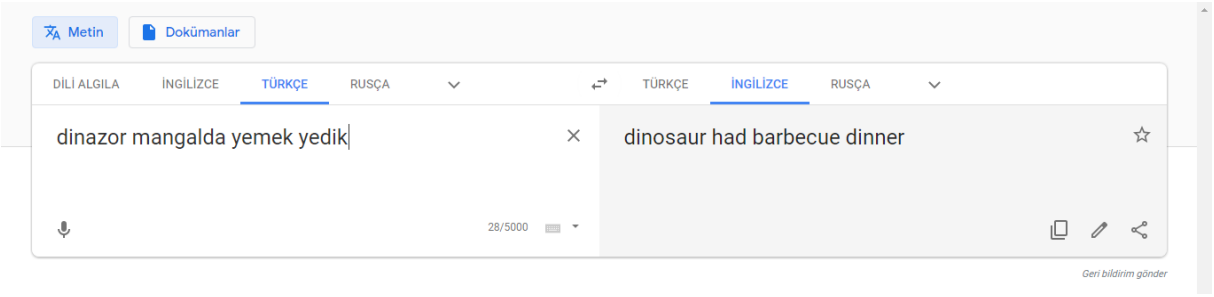


Figure 3. Mistranslation due to inaccurate spelling on https://translate.google.com

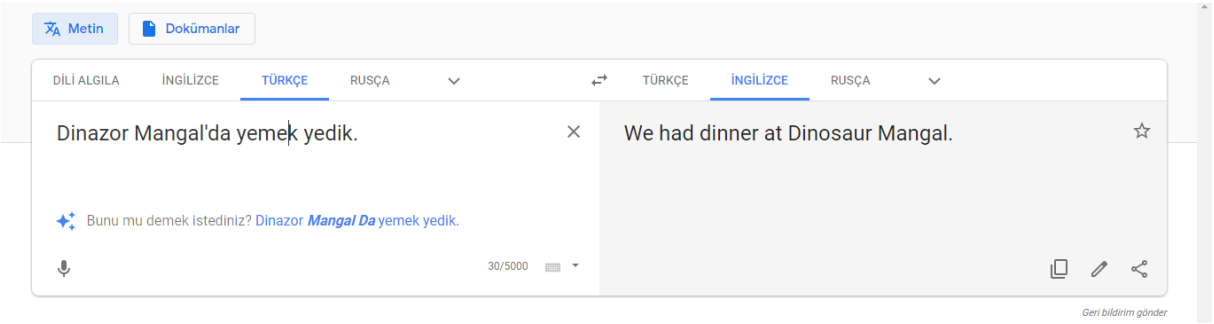


Figure 4. Output due to accurate spelling on https://translate.google.com

After the training session, the writing activities lasted for seven weeks; however, after the sixth week the students went on a one-week semester holiday.

3.7.2.2. Week 1. The objective of the writing exercise was "Students will be able to describe different environments in simple sentences and phrases" (MoNE, 2018, p. 24).

Aim: The students were asked to describe their dream room including examples of unusual uses of a clock on a paper.

Procedure: The students studied the grammatical structures of "have got/has got" and "there is / are" and the negative forms of them in addition to the prepositions of place. First

they read a small and simple text about a teenager's room and after some post-reading exercises they were given that instruction as a pre-writing activity: Think about a clock. In what other ways can you use it except for checking the time? Take small notes in 5 minutes.

After they were given 5 minutes to think of unusual uses of a clock, they were asked: Write at least ten sentences about your dream room and include at least one example of unusual uses of a clock in your dream room. Write in English and you can get help from the passage that we read before. You have fifteen minutes.

The researcher walked around the class during this period just to observe the students. She did not give any feedback on their writing in this step.

After they wrote their text in English, they got their mobile phones and translated their writing into Turkish sentence by sentence via Google Translate in order to check whether the responses were in consistent with their own sentences. If so, they put a tick; if not they edited their source sentences. While they were translating and editing their writing, the researcher walked around and this time she gave corrective feedback, especially when there was a mistake in their final editions in order not to result in fossilization (See Appendix 6).

3.7.2.3. Week 2. The objectives were "Students will be able to ask and answer questions about location of things and places" and "Students will be able to fill in a chart comparing cities in different countries/Turkey" (MoNE, 2018, p. 24).

Aim: The aim of the activity was to help the students create as many questions as possible without giving answers to them.

Materials: A chart of information about population, area, architecture and climate in Hong Kong and Vaduz.

Procedure: The students were provided the chart and they were asked to write as many questions as they could about the chart in fifteen minutes in English. They were encouraged to compare these two cities. They were not given any feedback during this process. Then they

translated their texts into Turkish via Google Translate and compared their own sentences to the responses by GT in order to detect if there were any differences in between. Lastly, they edited their sentences if needed. They got corrective feedback in this part (See Appendix 7).

3.7.2.4. Week 3. The third week was a revision week for the past studies so the objectives were a combination of four skills (MoNE, 2018, p. 24):

- Students will be able to introduce themselves and their family members.
- Students will be able to write simple sentences and phrases (a postcard, an e-mail or a hotel registration form).
- Students will be able to fill in a chart comparing cities in different countries/Turkey.
- Students will be able to describe different environments in simple sentences and phrases.

Aim: The aim of the activity was to make a revision of the two units they had studied before. The students were asked to write a mail to one of their friends about themselves, their hometown Bursa, their neighbourhood, their home and lastly their room. This time they were given forty minutes as the task required them to get help from the previous passages they studied or they needed pair work to write such a long and dense text.

The overall aim of the activity was to help the students flourish their elaboration skills in a piece of writing. In the previous reading and speaking activities, the students were taught about the importance of a topic sentence, conjunctions and phrases in organising a text. Therefore, students were asked to include topic sentences for each paragraph, some phrases such as “last but not the least”, “hope to hear from you soon”, “my nearest and dearest” or “hustle and bustle” and lastly some conjunctions –and, but, because-. These phrases were given as an example because the students studied these phrases during course hours.

Materials: No previously designed materials were used in this activity. They needed a piece of paper, a pencil, their course book, their mobile phones and the Internet.

Procedure: In the very beginning of the lesson, the students joined in a brainstorming about what they had learnt in their English classes up to then in order to heat up their memory. After the brainstorming, they prepared a mind map about what –among the items which arose in the brainstorming activity- they could include in a text about themselves. By means of this mind map, every student organized a four-paragraphed letter about themselves and decided what to write about in every paragraph. They were reminded to include the above mentioned phrases, conjunctions, topic sentences and a greeting sentence in their letters. They were given thirty minutes and they started to write in English. However, on seeing that they had difficulty in writing topic sentences, the instructor helped those in need to write one. Otherwise, they were discouraged, demotivated and were about to give up.

After they finished writing in thirty minutes, it was time for Google Translate. They got their mobile phones and translated their text sentence by sentence to check if they could give the intended meaning. If so, they put a tick. If not, they edited the sentence by typing in the Turkish sentence and getting the English translation on Google Translate. Lastly, the instructor gave corrective feedback if necessary to avoid fossilization (see Appendix 8).

3.7.2.5. Week 4. The objective of the fourth week was "Students will be able to ask and answer about their personal belongings" (MoNE, 2018, p. 24). It was the exam week in the school so the students could not focus on the lesson and they were not as motivated as the other weeks. So the aim was kept simple and the procedure was not as demanding as the other weeks.

Aim: The aim of the activity was to let students write about one of their favourite belongings and add some features in order to improve it.

Procedure: First of all, the students were asked about their most favourite belonging, excluding the technological ones. Then, a 'find someone who' activity started. They walked around the classroom to find someone who had the same favourite belonging. The matches sat

down together and started a group work. The outliners joined the group which appealed them most. The groups drew a picture of the object and were given five minutes to negotiate about how they could improve that object to make it more attractive. After five minutes the groups started to write as many sentences as possible using have got/has got. For example, 'My favourite belonging is a watch. It has got colourful numbers.' They pre-edited their writings with Google Translate (see Appendix 9).

Finally, the improved objects were presented to class by one of the members from each group and the class ranked them.

3.7.2.6. Week 5. The objective of the fifth week was "Students will be able to write a dialogue in order to invite a friend to cinema." (MoNe, 2018, p.24). It was the week that they started the new unit on their book *Teenwise 9*. The name of the unit was 'Movies' and the general objectives were how to make an invitation, accepting an invitation and refusing an invitation.

Aim: The activity had two main aims: flourishing the students' fluency and originality skills by means of Google Translate. First, the students studied a dialogue on their course books (Teenwise 9, p. 39) and revised how to make invitations, accept and refuse an invitation. In the sample dialogue, one of the speakers invited his friend to a movie; however, the other made an excuse and refused him. The students were asked to write a similar dialogue but this time they were supposed to make original excuses as many as possible in twenty minutes using the phrases they revised in the previous reading exercise.

Procedure: First of all, the students studied a dialogue from their textbook.

Vincent: Hi, Doreen. There is a new Batman movie on at the cinemas. How about seeing it on Tuesday afternoon?

Doreen: Well, I'd love to, but I can't. I have a photography class.

Vincent: OK. Let me check my schedule. I'm free on December 5th. Why don't we meet on that day?

Doreen: Sorry, but I can't make it. It's my best friend's birthday. Shall we meet on Thursday?

Vincent: It sounds great. Let's meet at the cinema at 5:30 (Teenwise 9, 2019, p. 39).

On reading the dialogue, the students sorted out the phrases into three groups: making an invitation, accepting an invitation and refusing an invitation. Then, they were asked to write a similar dialogue, but this time making original excuses as many as possible in twenty minutes in order to imply that Doreen did not actually want to join Vincent. The instructor walked around the classroom in order to encourage students to find original excuses without helping them with their writing. Finally, the students translated their English into Turkish via Google Translate to see whether they meant whatever they intended to mean. If not, they translated their sentences from Turkish to English and edited the sentences to write a final one (see Appendix 10).

After they finished editing, the instructor checked the sentences again in order to avoid any possible errors in the dialogues.

3.7.2.7. Week 6. On the sixth week the students continued studying the third unit, the name of which was Movies. The writing objective of the unit was "Students will be able to write their opinions on a blog" (pp. 23-25).

Aim: Among the tasks in Torrance Test of Creativity, 'just suppose' is a prompt to lead the participants write about the consequences of an unusual situation, for example just suppose that the clouds were attached with strings. What would be the consequences? The overall aim of this task is to determine fluency, flexibility, originality and elaboration. In the present study, the participants were beginners and they were not proficient enough to write

about such an unusual situation; therefore, the just suppose task was adapted in accordance with the writing objective provided by the MoNE.

Procedure: As a pre-writing activity they read three simple blog writings about three different movies and they studied some active vocabulary: cast, plot, setting, acting, to take place, I think, in my opinion, to me and then they filled in a table about the movie 'The Martian' (Teenwise 9, p.42).

After that, the students were given that instruction: Just suppose that you are a director and you shot a movie. Give it a name and write about its type, cast, the lead roles, setting and plot in twenty minutes. They filled in a table similar to the one about the Martian (see Appendix 11). Then they translated their sentences into Turkish via Google Translate and checked if they can convey the intended meaning. If there were any problems they entered the Turkish sentence and got the English version in order to write a final version. Lastly, the instructor gave feedback and corrected if there were any mistakes.

3.7.2.8 Week 7. The last week of the instruction process was the second step of the previous week. The objective of this week was the same: "Students will be able to write their opinions on a blog" (Teenwise 9, pp. 23-25).

Aim: The students had written about a 'just suppose' situation and had prepared a table about their imaginary movies. On the seventh week, four volunteers presented their movies, which were different from each other in terms of type and plot. Then, the students wrote a paragraph about their opinions choosing one of the presented movies. The aim of the activity was raising their originality in expressing their opinion.

Process: The students listened to four of their friends and picked one of the movies. They were given these questions as prompts in their writing:

- What is the name of the movie?
- What type of a movie is it?

- Who is in it?
- Who are the lead characters?
- Where does the movie take place?
- When does the movie take place?
- What is the movie about?
- What do you think of the movie?

They were given twenty minutes to write in English and then they translated their blog writing into Turkish via Google Translate. If there were any problems, they reversed the languages and got an English version from the Google Translate as a base to edit his/her sentence (see Appendix 12).

3.8. Data Analysis Procedure

There were two sets of data in the present study: pre and post-test analysis as quantitative data and semi-structured interview recordings as qualitative data.

In detail, quantitative data were gathered from the written products which were scored in terms of three aspects: overall creativity by five raters in terms of the CAT; fluency, flexibility, originality and elaboration by the researcher; grammar and lexical errors by the researcher. The scores were analysed by using the Statistical Package for Social Sciences (SPSS), version 22.

3.8.1. Creativity assessment. Assessing creativity is a multi-layered issue in terms of reliability and validity. Some standardized tests have been generated by many researchers, such as “the Alternate Uses” test by J. P. Guilford in the late twentieth century, “the Torrance Test of Creativity” (TTCT; Torrance, 1974), Abbreviated Torrance Test for Adults (ATTA; Goff, 2002) and the Abedi Test of Creativity (ATC; Abedi, 2000). However, these tests are addressing the creativity in general, different from the aim of this present research which is to assess the creativity in students’ written products in EFL. What is more the assessment of

these tests requires creativity professionals. However, the aim of this research is to find out the impact of using Google Translate in writing activities on the creativity of participants' written products.

Therefore another way of assessing creativity is adopted in the research: the Consensual Assessment Technique (CAT; Amabile, 1982). The CAT helps the researcher to solely focus on the creativity in just the written products of the participants and make an in-group rank order. Hennessey and Amabile (1999, p. 347-359) define the CAT as "a technique used for the assessment of creativity and other aspects of products, relying on the independent subjective judgements of individuals familiar with the domain in which the products were made". Namely, in the discourse of the present research, English teachers who are familiar with teaching writing to beginners, made subjective judgements on the written products of the students and assessed the creativity in these products independent from each other.

In this research, five English teachers assessed participants' products which they wrote as pre and post-tests. Each of the raters graduated from the English Language Teaching department and had been teaching English for at least ten years at state schools. The raters were not provided any background knowledge about the participants or the research, neither about creativity assessment. They were asked to make an in-group ranking and to ignore grammatical and lexical errors as the participants were beginner learners. In terms of the CAT, the raters were given the following instruction, which was adopted from Baer's study:

> There is only one criterion in rating these products: creativity. I realize that creativity doesn't exist in a vacuum, and to some extent creativity probably overlaps other criteria one might apply — aesthetic appeal, organization, richness of imagery, sophistication of expression, novelty of word choice, appropriateness of word choice, and possibly even correctness of grammar, for example — but I ask you to rate the products of our ninth graders solely on the basis of your thoughtful – but – subjective

> opinions of their creativity. The point is, you are the expert, and you needn't defend your choices or articulate a definition of creativity. What creativity means to you can remain a mystery — what I want you to do is use that mysterious expert sense to rate the writings for creativity (Baer, 1994, pp. 39–40).

Then, the raters studied on the artefacts independently and first piled them into three (unsatisfactory, creative and high creative). Then they subdivided each group into two and got 6 levels of creativity (Baer et al., 2004). The raters did not discuss about their judgements until after the rankings were submitted to the researcher.

The researcher herself did not rank the products in terms of the CAT because in addition to the CAT, she analysed the artefacts both in pre and post-tests in terms of four indices of creativity: fluency, flexibility, originality and elaboration. As it is discussed in Runco and Acar (2012), several ways of raising reliability and validity in creativity studies have been suggested by many researchers, such as simply adding the scores or getting a proportional score or scoring only fluency and originality or using median weighs. Therefore, the researcher scored the four indices independently and did not get a total score of creativity in order to get rid of such statistical biases. In addition, it is cited in Cramond et al. (2005, p.284) "Torrance maintained that the composite score was not the most useful way to look at a person's creative functioning because he knew it could mask individual strengths". As for fluency, every sentence which was grammatically accurate enough to get the supposed meaning was simply counted. The flexibility scores were the number of the ideational categories that were generated by each participant, such as physical appearance, personality, daily routines or abilities. A dichotomous scoring was applied for originality and elaboration. Every idea which was unique in the group got 1 point in order to find out the originality scores. As for the elaboration scoring, the artefacts in pre-test were analysed and three ways

of elaboration were determined: giving a title, expressions to emphasize an idea and conjunctions. Every elaboration element got 1 point.

3.8.2. Grammatical and lexical error analysis. In order to give an answer to the second research question, the products of the participants were analysed in terms of grammatical and lexical errors. One of the aims of the research is to investigate the impact of using Google Translate in writing activities on their grammar and lexical knowledge. In order to grade their writings in terms of grammar, the Surface Structure Taxonomy (Dulay et al., 1982) was implemented. The taxonomy groups the grammatical errors into four categories: omission, addition, misformation and misordering. Every piece of writing was analysed and the errors were detected and labelled as 'om' for omission, 'ad' for addition, 'mf' for misformation and 'mo' for misordering. Then they were simply counted, 1 point for each error to find out a total score for each category (See Appendix 13).

As for the lexical knowledge, the idea by Barkaoui and Knouzi (2012, p. 93) was adapted. They suggest that there are two categories in lexical errors: "lexical choice errors and lexical form errors". If the participant chose a wrong word, it was regarded as lexical choice error. For example, handsome for beautiful 'she is handsome' or everyone for everyday 'he cries everyone'. The other type of lexical errors is lexical form errors. For example, he for his 'he name is Hasan' or has for have 'I has a dog'. However, this type of lexical errors are very much like the same as the misformation grammar errors. That is why this study did not make a distinction between two types of errors and the related errors were counted as misformation errors (See Appendix 13).

To sum up, the written products were scored in terms of three aspects: overall creativity by five raters in terms of the CAT; fluency, flexibility, originality and elaboration by the researcher herself; grammar and lexical errors by the researcher herself. The scores were analysed by using the Statistical Package for Social Sciences (SPSS), version 22.

Chapter 4

Results

4.1. Introduction

In this chapter, the quantitative and the qualitative results will be presented in order to give answers to every research question separately.

4.2. Results of the Quantitative Data Set

The quantitative data set includes data about creativity of the products and the students' grammatical and lexical errors in order to seek answers for the first and the second research questions. In the analysis of the data, the researcher was assisted by Assoc. Prof. Dr. Deniz Sığırlı, who teaches in the department of Biostatistics in T.C. Bursa Uludag University Medicine Faculty.

4.2.1. Creativity. The aim of this part is to determine the impact of using Google Translate on the creativity in writing products of the participants. The related research question is:

RQ 1: How does implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, elaboration and originality?

It is important to note that the study has a one-group experimental design, which means the data were gathered within the same group with pre and post-tests. In order to seek for an answer to this research question a two-paced analysis was held. The first one was to examine their overall creativity so the CAT was applied and the other one was to determine their fluency, flexibility, elaboration and originality scores independently.

4.2.1.1 The Consensual Assessment Technique. This technique has long been used in the creativity research field since Amabile (1982) first introduced it by suggesting that a consensus among experts in the same domain as the creative products could detect the level of creativity. The experts decide about the creativity of the products in relation to one and other

within a group of products and they are not asked to defend their perception of creativity (Baer & McKool, 2009).

In the present research, five randomly selected English teachers evaluated both the pre-tests and post-tests independently. They were not in relation with each other during the assessment process in order to avoid the possible biases as suggested by Baer and McKool (2009). What is important in the CAT is to be able to provide an agreement among the raters and it can be proven with a satisfactory level of inter-rater reliability. Many studies that conducted the CAT have yielded satisfactory inter-rater reliability scores. For instance, in Baer's study in 2003, there were three groups that were assessed by independent raters with acceptable inter-rater reliabilities, the coefficient alpha of which were .83, .79 and .77 (Baer, 2003). Another study with 13 raters by Baer et al. (2004) reported .66, .76 and .79 coefficient alphas. Phonethibsavads, Bender and Peppler (2019) report a high agreement among the raters with .81 coefficient alpha.

The inter-rater reliability of the present research was calculated with an intra-class correlation coefficient test and it was acceptable with a coefficient alpha of .86 for the pre-test, and of .85 for the post-test. A coefficient higher than .80 is labelled as "almost perfect" by Landis & Koch (1977). As it is cited in Baer et al. (2004), Amabile, who is the pioneer in this technique, agrees that correlation coefficient between .70 and .80 shows a strong agreement among the raters, which is already lower than the coefficient alphas of the present study. In other words, the agreement among the raters in the present study can be labelled as high with an almost perfect degree.

Apart from getting a high degree of agreement among the raters, the descriptive statistics of a Shapiro-Wilk test indicated that the data gathered from the raters were normally distributed both for pre and post-tests. Thus, a paired-sampled t-test was applied in order to

determine if there were any statistically significant difference between the pre and post-tests and the mean ± standard deviation values were presented in the Table 2.

Table 2

Descriptive statistics of the raters for pre and post-tests (N=5)

	N	Mean	Std. Dev.	p
Pre-test	5	12.11	4.68	
Post-test	5	13.37	5.46	.19

As seen in the Table 2, the mean value of the raters in the pre-test was calculated as M = 12. 11 (SD = 4. 68) and with a slight increase in the post-test mean value was calculated as M = 13. 37 (SD = 5. 46). The difference between the pre and post-tests' mean values was 1. 26, which did not yield a statistically significant difference.

The CAT provided valuable information about the participants when the descriptive statistics were examined in detail for every student, who were labelled randomly with ordinal numbers. For instance, the most creative five products in the pre-test belonged to relatively the 1st, 15th, 23rd, 30th and 4th students while the most creative 5 products in the post-test were relatively the 10th, 7th, 3rd, 9th and 1st students. As for an individual comparison in students' pre and post-test products, the largest positive difference between the pre and the post-test artefacts belonged to the 10th, 7th, 16th, 34th and 9th students. On the other hand, the largest negative difference was of the 1st, 28th, 22nd, 5th and the 31st students.

4.2.1.2. Components of creativity in pre and post-tests. Creativity is a multi-layered issue and it has long been searched from different aspects. In a number of studies conducted by Guilford, one of the pioneers in the field, four components of creativity were identified by means of factor analysis: fluency, flexibility, originality and elaboration (Hickey, 2001; Kim, 2006). As discussed in the chapter 2, fluency is the total number of generated ideas; flexibility

is the total number of ideational categories; originality is the total number of the unique ideas among the participants; and elaboration is the total number of the details which elaborate the product. In order to detect these numbers, the researcher herself examined the products of the participants in pre and post-tests one by one in detail and simply got a total number for all of these components.

4.2.1.2.1. Fluency in the pre and post-tests. In order to find out the fluency values of the participants every meaningful sentence of each artefact was simply counted ignoring the simple grammatical or lexical errors. When the results were tested through a Shapiro-Wilk test, it was found out that the homogeneity of the variances was not normally distributed. Therefore, median scores (min-max) were tabled in Table 3.

Table 3

Descriptive statistics for the fluency scores in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	10 (4 – 29)	
Post-test	35	15 (7 – 28)	< .001

As seen in the Table 3, while the median value (min – max) for the pre-test was 10 (4 – 29), it was 15 (7 – 28) for the post-test. As the homogeneity of the variances were not normally distributed, a Wilcoxon test was applied to see the significance level of the difference between the two tests and it was found out that there were a statistically significant difference in the fluency scores of the pre and post-test products.

4.2.1.2.2. Flexibility in the pre and post-tests. The researcher analysed the written products one by one and categorised the ideational groups in them, for example age, ability, job, background knowledge, feeling, family, friends. While the student who presented the largest amount of ideational category in his writing had ten categories in the pre-test, it was

thirteen in the post-test. These results were tested through a Shapiro-Wilk test and found out to be normally distributed, thus mean scores and standard deviation were tabled in the Table 4 below.

Table 4

Descriptive statistics for the flexibility scores in pre and –post-tests

	N	Mean	Std. Dev.	p
Pre-test	35	5.88	1.95	
Post-test	35	8.82	1.79	< .001

As seen in the Table 4, the mean value of the participants in the pre-test was calculated as M = 5.88 (SD = 1.95) and with an increase in the post-test mean value was calculated as M = 8.82 (SD = 1.79). The difference between the pre and post-tests' mean values was found to be $p < .001$ with the help of a paired samples T-test, which yielded a statistically significant difference.

4.2.1.2.3. Originality in the pre and post-tests. Originality is one of the main aspects of creative products. As Runco and Jaeger suggest "Originality is undoubtedly required. It is often labelled novelty, but whatever the label, if something is not unusual, novel, or unique, it is commonplace, mundane, or conventional. It is not original, and therefore not creative" (Runco & Jaeger, 2012, p. 92). Taking this into consideration, the researcher analysed the products in terms of original, in other words unique, ideas. Every idea was considered and those which were not written by anyone else in the group was regarded and counted as original. After the analysis, the homogeneity of this variance was found to be not normally distributed by means of a Shapiro-Wilk test. Hence, the median scores (min-max) were presented in the Table 5.

The findings suggested that the median score (min-max) of the pre-test was 2 (0-10) and it was 9 (2 – 27) in the post-test. There was a 7-point difference between the pre and the post-test. When this difference was tested through a Wilcoxon test – as they were not normally distributed – it was found out that the increase between the pre and the post-tests were statistically significant with a p value lower than .001.

Table 5

Descriptive statistics for the originality scores in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	2 (0 – 10)	
Post-test	35	9 (2 – 27)	< .001

4.2.1.2.4. Elaboration in the pre and post-tests. Elaboration is described by Guilford (1966, p. 188) as "finishing touches". It can be anything that embroiders the product. When the written products were analysed by the researcher, it was found out that the students tried to elaborate their writing with a title, some adverbs or adjectives to give details, some exclamation expressions, conjunctions and sequencing words. When they were simply counted, an elaboration score was determined for each student. The homogeneity of this variance was found out not to be normally distributed by means of a Shapiro-Wilk test so the median scores (min-max) were tabled in the Table 6.

Table 6

Descriptive statistics for the elaboration scores in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	1 (0 – 4)	
Post-test	35	3 (0 – 13)	< .001

As it can be seen in the Table 6, the median (min-max) of the pre-test was 1 (0 – 4), it was 3 (0 – 13) in the post-test. It is clear that there was an increase in the number of the elaboration items in the students' writings. While the largest amount was 4 in the pre-test, it was 13 in the post-test. This difference was calculated with a Wilcoxon test and the difference was found to be statistically significant.

4.2.2. Grammatical and lexical knowledge. The aim of this part is to determine the impact of using Google Translate in writing activities on the grammatical and lexical knowledge of the participants. The related research question is:

RQ 2: How does implementing machine translation in writing activities in English affect the participants' grammatical and lexical knowledge?

In order to seek an answer to this questions, the grammatical and lexical errors were coded in both their pre and post-tests. The grammatical errors were categorized under four titles: omission, addition, misformation and misordering as Dulay et al. suggested in their book Second Language (1982). The lexical errors were going to be codified under two titles: form and lexical choice; however, as the misformation errors and the form errors cannot be distinguished from each other, only the lexical choice errors were counted.

4.2.2.1. Omission errors. These types of errors are about missing elements that should be in a grammatically accurate sentence. The learner may forget to add a helping verb or a suffix although they are necessary. Although it may sometimes result in meaningless or ambiguous sentences, these are mostly slight errors.

The results of the pre and post-tests in the present research were tested via a Shapiro-Wilk test and because this variance was not normally distributed, the median (min-max) values were given in the Table 7. As it can be seen from the Table 7, the median scores of pre and –post tests were the same, namely 6, although the maximum value of the post-test slightly increased. The difference was calculated by means of a Wilcoxon test because the variance

was not normally distributed and the result did not yield a statistically significant difference with a p value of .09.

Table 7

Descriptive statistics for the omission errors in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	6 (0 – 28)	
Post-test	35	6 (1 – 37)	.09

4.2.2.2. Addition errors. These errors are the existence of an extra element which should not appear in a grammatically accurate utterance. The learner may utilize more than one helping verb as in 'I am do crazy' or may keep plural –s although it is not necessary as in 'childrens'. The products were analysed in terms of addition errors and the findings were found to be not normally distributed with the help of a Shapiro-Wilk test. Therefore the median (min-max) scores were given in the Table 8:

Table 8

Descriptive statistics for the addition errors in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	2 (0 – 9)	
Post-test	35	2 (0 – 22)	.71

The Table 8 demonstrates the descriptive statistics of addition errors in the pre and post-tests. The median scores (min-max) were the same in both tests, thus the p value was .71. It did not show a statistically significant difference.

4.2.2.3. Misformation errors. As the name suggests, this type of errors are about the accurate formation of words. The learner may not use the correct form of the helping verb, as

in ‘they doesn’t’ or he/she may not decide on the correct form of a possessive adjective as in ‘he name is Hasan’. After analysing and counting the misformation errors in the participants’ writings, the variance was tested via a Shapiro-Wilk test and the homogeneity of it was found to be not normally distributed. Therefore, the median (min-max) values were demonstrated in the Table 9 below:

Table 9

Descriptive statistics for the misformation errors in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	1 (0 – 6)	
Post-test	35	0 (0 – 9)	.05

As the Table 9 suggests the median value (min-max) in pre-test was 1 (0-6) and it was 0 (0-9) in post-test. The difference between these values was examined through a Wilcoxon test and it was found out that there was not a statistically significant difference.

4.2.2.4. Misordering. The learners sometimes cannot decide on an accurate sequence of the utterances, these errors are called misordering. For example, he eats always cucumbers for breakfast. This type of errors were detected and counted by the researcher and the results were tested with a Shapiro-Wilk test. The variance was found out to be not normally distributed so the median values (min-max) were tabled in Table 10:

Table 10

Descriptive statistics for the misordering errors in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	0 (0 – 4)	
Post-test	35	0 (0 – 3)	.68

As the Table 10 suggests the median value (min-max) in pre-test was 0 (0-4) and it was 0 (0-3) in post-test. The difference between these values was examined through a Wilcoxon test and it was found out that there was not a statistically significant difference.

4.2.2.5. Lexical choice errors. Language learners sometimes have difficulty in choosing the right word. For example, he cries everyone. Actually, the intended meaning is 'he cries every day'. However, for some reasons the learner may make mistakes about lexical choices. Such kind of errors were detected and counted in the writings and it was found out that the variance was not normally distributed with a Shapiro-Wilk test so the median (min-max) values were tabled in the Table 11.

Table 11

Descriptive statistics for the lexical choice errors in pre and –post-tests

	N	Median (min-max)	p
Pre-test	35	1 (0 – 5)	
Post-test	35	1 (0 – 14)	.69

The findings in the Table 11 show that the median value (min-max) in the pre-test was 1 (0-5) and it was 1 (0-14) in the post-test, which actually does not yield a statistically significant difference with a p value of .69.

4.3. Results of the Qualitative Data Set

In this part of the Results chapter, the qualitative data obtained from the semi-structured interview are presented. These interviews were conducted so as to give an answer to the third question of the research, which is:

RQ3: What are the perceptions of high school EFL students on the use of machine translation in writing activities in English classes?

It is important to note that the interviews were in Turkish as the participants were beginner learners. Among 35 participants, six students (3 females, 3 males) gave a consent for making the interview and they were asked questions about the impact of using Google Translate in their writing activities. The interviewees were given pseudo names. The interview was recorded and interpreted by the researcher and the results were presented in line with the interview questions. It is noteworthy to mention that the interviews were conducted one week after the post-test and they were semi-structured. Only five questions were determined beforehand, however some more questions were generated during the interviews in order to get a deeper understanding of the participants' perceptions. The participants' answers to these questions were coded into different categories and presented in the Tables 12, 13, 14, 15 and 16.

4.3.1. Question 1: What were the advantages and disadvantages of implementing machine translation in writing activities?. The interviewees gave more or less similar answers to this question (Table 12). Deva, Eser and Rahman thought that it was an advantage to get a translation of everything from a single word to a whole paragraph by the help of Google Translate in a fast and easy way just in a second. Deva was also pleased with the fact that GT offered translation in a number of languages. Another feature which was mentioned by Eser was that the application was all free and it provided a priceless help in their writing. Kaan was in favour of write, speak and snap modes of GT, which relatively enables the users to get a translation by writing directly on the screen, by speaking in the microphone and by taking a picture of the text. Kaan also added that it enhanced their vocabulary knowledge.

Yeşim and Deva expressed something very important. They thought that the instruction process provided them with high marks. They were of the opinion that the utilisation of Google Translate in the writing activities brought them an academic success.

There was a disagreement among the interviewees about the accuracy of GT. Although two of the interviewees were of the opinion that GT sometimes provided them a chance to correct their grammatically wrong sentences and helped them with pronunciation and spelling as well, all of them admitted that it sometimes translated inaccurately. The reason for this inaccurate translations was explained by Kaan as their own fault, because he thought GT was sensitive about the spelling mistakes and even a missing comma or full stop might result in different translations (see Appendix 14). Another explanation about this issue was by Deva. She was of the opinion that GT was generally useless without an Internet connection although it offered offline translation. She preferred using the website to the application as it was better at translation than the application.

Table 12

Summary of the categories, subcategories and the corresponding codes emerging from the first question in the semi-structured interviews

	Category	Subcategory/ Frequency	Code	Example Meaning Unit
Q 1	Advantage	Features of Application (n=3)	free	"It is an advantage that the application is free (Kaan). Its positive feature is that it is free (Sezer)."
			fast	"I can find an unknown word fast (Rahman)."
			accessible	"You can have access with a Wi-Fi (Eser)."
		Lexical Knowledge (n=4)	Pronunciation, Spelling, Use	"I can learn the pronunciation, spelling and use of a word that I look up (Kaan). It translates the words that I do not know (Eser). It helped me with the unknown words (Deva). It helped me a lot with word translations (Ezel). It helped me express myself better and in an easy way because I had difficulty with my vocabulary knowledge (Deva)."
		Grammar knowledge	Word order	"I translated sentences when I did not know the grammar rule (Rahman). I wrote the Turkish version and got a translation if I did not know the grammar (Eser). I have a problem with

	(n=3)		forming sentences but I could make sentences easily with the help of GT (Ezel). It helps us start the sentence (Ezel)."
	Academic achievement (n=2)		"I wrote in accordance with what we had studied in that process in the writing exam. I could not get higher marks than 60 in the secondary school but I got an 89 in the exam (Yeşim). I was not good at English in secondary school but I got an 85 in the exam this year (Deva)."
	Further studies (n=2)	encouragement	"I was afraid of learning a new language before but now feel encouraged to learn even other foreign languages. I want to study language in the eleventh and the twelfth grade (Yeşim). I want to write a book in English (Deva)."
Disadvantage	Inaccurate Translation (n=6)	Internet access	"It may give inaccurate translations offline (Kaan / Deva)."
		Punctuation	"It may translate inaccurately if we make punctuation mistakes (Kaan)."
		Inaccuracy	"It can sometimes translate inaccurately (Yeşim / Eser / Ezel)."
		Inability to get nuances	"It can sometimes translate inaccurately because it cannot get the nuances as well as a human (Rahman)."

4.3.2. Question 2: Do you think implementing machine translation affected your creativity in writing activities?. There was a controversy among the interviewees about this question (Table 13). While most of them agreed that it affected their creativity positively, one of them thought that it had nothing to do with it. Deva thought that she had difficulty in expressing her ideas in English as a low-proficient learner, and GT helped her when she was in need of some vocabulary and grammatical patterns. Furthermore, she felt like GT helped her to generate new ideas and express them accurately in English. She even admitted that she felt so confidence in GT that she planned to write a book in English as she enjoyed reading and writing.

On the other hand, Kaan associated creativity with a powerful imagination skill and he assumed that GT could not help flourish their imagination. Nevertheless, the rest admitted that they put their imagination into words in English easily with the help of GT.

Table 13

Summary of the categories, subcategories and the corresponding codes emerging from the second question in the semi-structured interviews

	Category	Sub-category/ Frequency	Code	Example meaning unit
Q 2	Impact on creativity	No impact (n=1)	Imagination	"If a person has a powerful imagination, GT has nothing to do with it (Kaan)."
		Positive impact (n=5)	Assistance to express creativity	"I sometimes could not find a proper word but GT helped me find it and write more sentences (Eser). I could more easily express what I imagined (Rahman). It helped with the word choice and word order while I tried to express myself (Yeşim). I could not make sentences before GT but now I can while expressing my thoughts (Ezel) It is easier with GT to express my thoughts because my grammar and lexical knowledge is not sufficient (Deva)."

4.3.3. Question 3: Do you feel like you learnt new things about English while implementing machine translation in writing activities?. The interviewees shared almost similar ideas about this question (Table 14). The most commonly stated idea was that GT thought them new vocabulary. Kaan was in favour of the fact that it enhanced their vocabulary knowledge along with its correct pronunciation and spelling. Deva stated something important that she took an interest in learning new vocabulary with GT. Eser expressed that he was very bad at making sentences in English and GT helped him learn new things about grammar and vocabulary.

Apart from grammar and vocabulary, Deva mentioned that GT encouraged her for written communication with friends from different countries and she could figure out the reasons behind the errors of GT.

Table 14

Summary of the categories, subcategories and the corresponding codes emerging from the third question in the semi-structured interviews

	Category	Sub-category/ Frequency	Code	Example meaning unit
Q 3	Educational tool	Grammar knowledge (n=5)	Word order	"I learnt how to order the words to make sentences (Rahman). I learnt a lot of new things, grammar, word order (Eser). I learnt new things about word order (Yeşim). Now I can make sentences and I can now order the words correctly to make sentences (Deva)."
			Pre-fix / Suffix	"It help me with some suffixes, for example I always forget to put an –s in the Present Simple Tense (Ezel)."
			Grammatical awareness	"I can understand the reason behind the GT's mistakes (Deva)."
		Lexical knowledge (n=4)	Spelling Pronunciation Use	"I learnt something about new words, their spelling, and pronunciation and how to use them in sentences (Kaan). I learnt new vocabulary (Eser). It makes it easy to memorise

new vocabulary and their pronunciation (Yeşim). It enhanced my vocabulary knowledge (Deva).”

4.3.4. Question 4: Did you experienced any problems while using machine translation?. Among the interviews only one (Rahman) did not mention about any problems while using GT; however, the rest of the participants stated that they got inaccurate translations if they were offline (See Table 15).

Table 15

Summary of the categories, subcategories and the corresponding codes emerging from the forth question in the semi-structured interviews

	Category	Sub-category/ Frequency	Code	Example meaning unit
Q 4	Technical problems	None (n=1)		"I did not face any problems (Eser/ Rahman)."
		Inaccurate translation (n=4)	Internet access	"If it is offline, it gave wrong translations (Kaan / Yeşim / Ezel / Deva)."

4.3.5. Question 5: Had you used machine translation before we started this study?. The answers to that question revealed two reasons to use GT: for classes and for private concerns (see Table 16). 2 of the students reported that they had rarely utilised GT for classes, especially because of the inaccurate translation. However three of them reported that they had frequently used GT for their classes. Deva mentioned that she had got help for the exams. The other two mentioned they had used GT for their English homework. Yeşim said that her classmates and she had used GT even though their teacher had not permitted it. As for the private concerns, Eser, Yeşim and Deva applied GT in their daily life for different reasons. Eser got help from GT for his computer games. Yeşim wrote in different languages in order to express her feelings so that nobody could read what she wrote and also she tried to learn

German and Spanish. Deva likes writing and she translated her writings into English by means of GT. In addition to these answers, all of the interviewees mentioned that they wanted to use GT for their further studies in their language learning process.

Table 16

Summary of the categories, subcategories and the corresponding codes emerging from the fifth question in the semi-structured interviews

	Category	Sub-category/ Frequency	Code	Example meaning unit
Q 5	Background	Rare use (n=2)	For classes	"Rarely. After I had realised that it might lack accuracy, I did not address it again (Kaan). I once utilised GT for my English homework (Eser)."
		Frequent use (n=4)	For classes	"My friends and I used GT in secondary school for our homework although our teacher did not let us (Yeşim). I have always used GT for my English homework (Ezel). I got help from GT for the writing exam (Deva)."
			For private concerns	"I frequently use GT to translate unknown words while I am playing computer games (Eser). I usually try to write in a different language when I feel blue, so that others cannot read them. I use GT for this. I also tried to learn German and Spanish last summer and got help from GT (Yeşim)."

Chapter 5

Discussion

5.1. Introduction

This chapter presents the discussion of the results gathered from the analysis of the qualitative and quantitative data in line with the research questions. The research questions of the study are as follows:

RQ1: How does implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, elaboration and originality?

RQ2: How does implementing machine translation in writing activities in English affect the participants' grammatical and lexical knowledge?

RQ3: What are the perceptions of high school EFL students on the use of machine translation in writing activities in English classes?

In order to answer these questions the pre-test taken without any help from outside before the instruction process had started provided important information about the present creativity of the participants' written products; the seven-week instruction period helped the participants to get familiar with the strengths and weaknesses of this application. At the end of that period, the participants took a post-test and wrote directly in English with the help of Google Translate.

The quantitative results from the analysis of the products will be discussed in order to answer the first and the second questions in terms of creativity, grammatical and lexical errors to see if there was an impact of implementing Google Translate in their writing.

The comparison between the level of creativity of the students' products written in the pre and post-test addressed by the first research question will be discussed regarding the SPSS analysis of the data gathered by means of the Consensual Assessment Technique (the CAT; Amabile, 1982) and the scores of the products in terms of their fluency, flexibility, originality

and elaboration. Additionally, the products were analysed in terms of grammatical and lexical errors, addressed in the second research question. The statistical findings from this analysis will be discussed in detail in the chapter.

While the quantitative data put a light on the first and the second questions, the qualitative data provide insight for the third research question. The results from the semi-structured interview will be discussed in order to illuminate the question about the perceptions of the participants on machine translation, Google Translate for this study. The qualitative data was recorded and coded by the researcher herself after the semi-structured interviews.

5.2. Creativity in the Written Products

In the pre-test, the participants were not allowed to get help from Google Translate. They had thirty minutes to write anything they wanted about a stick-man. After a seven-week instruction, they wrote about a blurry image of a man staring into a puddle in twenty minutes, but this time they were allotted another twenty minutes to pre-edit their writing with Google Translate.

The creativity in the written products were analysed with two different techniques. The first one was the Consensual Assessment Technique, which provided an overall creativity score for each product and the second one was implied in order to get a denser analysis of fluency, flexibility, originality and elaboration in the products.

5.2.1. The Consensual Assessment Technique. The primary aim of the study is to find out whether creativity in low proficient English language learners' written products will be affected when they have an assistance with their limited grammar and lexical knowledge. For the present study, this assistant is the most frequently applied online machine translator, Google Translate. The present study tries to provide an inside about the impact of implementing Google Translate in writing activities on creativity.

As creativity is an intertwined issue, there is even a controversy even about the definition of it. While some researchers have focused on the common personal features of people who are known as creative, such as musicians, actors or authors, some others have approached the subject from a very different point of view and have focused on the features of the products (Kaygın & Çetinkaya, 2015). Teresa Amabile is one of those researchers who are concerned with the features of the products and she defines creativity as "the production of a novel and appropriate response, product, or solution to an open-ended task" (Amabile, 2012, p. 3) and "A product or response is creative to the extent that appropriate observers independently agree it is creative" (Amabile, 1982, p. 1001). Taken these definitions into consideration, the researcher searched for creativity in the written products of low proficient English learners by means of the Consensual Assessment Technique (the CAT; Amabile, 1982). And this effort was the core for the present study addressing the first research question, which is "How does implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, elaboration and originality?".

In order to compare the creativity of the products in pre and post-tests, five randomly chosen English teachers rated them independently. They were not given any prior instruction about creativity as suggested in the CAT. They were asked to pile the creativity in the products first in three levels - low creative, creative and high creative – according to their own subjective creativity definition. Secondly, they divided each group into two and presented the researcher a six-grouped rating. They ignored grammatical and lexical errors as far as the sentence conveyed a meaningful message as the participants were low proficient learners.

Providing an inter-rater reliability in the CAT is one of the most essential part of the technique. In a study by Cseh and Jeffries (2019), it is argued that some traditional techniques, such as Cronbach alpha or intra-class coefficient, are valid in order to calculate the inter-rater reliability in the CAT studies and they produce more or less similar results. In

the present research, the inter-rater reliability in the assessment was calculated with an intra-class correlation coefficient (ICC) and it was found to be .86 for the pre-test and .85 for the post-test. These results suggest that there is a strong agreement among the raters about assessing creativity.

These results are consistent with many of the studies implementing the CAT. For example, in a study by Hickey (2001), raters (N 61) used a seven-point Likert scale to judge the creativity of the musical products. The inter-rater reliability was calculated separately in raters groups: the lowest reliability score belonged to the composers with .04 and the highest score was of the teachers with .91. This study shows that the teachers provide more agreement about judging the creativity in the students' products.

In Baer et. al's (2004) study, the inter-rater reliability among the raters (N 14) was calculated as .94 for the stories and .95 for the personal narratives. These values are labelled as quite high levels by the authors. The raters were experienced judges about rating the secondary school students' writings. They were creative writers, editors, teachers and psychologists. The authors suggest that small numbers of judges will provide more valid results taking Amablie's validation studies in 1983 and 1996 into account.

In the aforementioned studies, the judges were experts in different occupations, such as teachers, psychologists, composers. Unfortunately, the present research is limited in that all the raters are teachers. What is more, a study by Gralewski and Karwowski (2018) argues that teachers are not good at detecting creativity potentials in their students. They assert a number of studies which show a low correlation between the teachers' nominations and the creativity potentials of the students such as Karwowski (2007) and Karwowski, Gralewski & Szumski (2015) and they suggest that "the accuracy of the teachers' creativity nominations is only slightly higher than a coin toss" (Gralewski & Karwowski, 2016, p. 157). The results of their experimental study revealed that one –third of the teachers in the study misjudged the features

of the students which were actually regarded as creative according to the creativity literature. Nevertheless, Amabile (1982) presents the requirements about the judges in the CAT and she suggests the raters have some experience within the domain and be familiar with the assessing issue. What is more in their book Essentials for Creativity Assessment Kaufman, Plucker and Baer (2008, p. 74) make a clear explanation about the qualifications of the judges:

> Judges should have some familiarity with the population from which the subjects are drawn. For example, judges of middle school student work should have some familiarity with middle school student productions in the domain in question. A Nobel Prize-Winning physicist might not be as appropriate as a CAT judge of the creativity of middle school science fair projects as college science professor who has worked with middle school students in the past (Kaufman et al., 2008, p. 74).

Considering all of these explanations in the literature, the fact that the judges in the present study have at least ten years of teaching experience well-suits the requirements. Moreover, the judges in the present study did not judge the creativity of the students, instead they judged the creativity in their products, which is not the concern of the Gralewski and Karwowski's (2018) study. What is more, the raters did not know anything about the participants as the products were not labelled with their names.

An outstanding comment about the inter-rater reliability in the CAT is that "inter-judge reliability in this method is equivalent to construct validity" (Amabile, 1982, p. 1002). It can be inferred from her argument that a product should be rendered unquestionably creative to the extent that it is consensually agreed as creative by a group of appropriate experts. Nevertheless, she makes a distinction between the judgements of aesthetic appeal and judgements of technical goodness (Amabile, 1982, p. 1002). She suggests to demonstrate the judge that it is reasonable to separate these two elements in order to avoid the raters from

rating a product as creative just because it appeals to aesthetical senses or it includes good technical features. This separation is also essential for the discriminant validity of the study.

Taking these into consideration, the present study can be said to have both construct and discriminant validity since the inter-rater reliability is quite high and provides a construct validity. As for the discriminant validity, the participants were not required to add any details to their writing such as drawings, colours or different types of papers. They just wrote on a plain paper with their lead pencils and drew nothing. This provided the raters with an opportunity of focusing plainly on the creativity in the written product. Besides, the raters were informed about not to consider the grammar or lexis errors unless they hindered them to get a meaningful message. By means of this, the raters stayed on the alert for creativity, not for the errors.

As for the number of the judges, it is suggested by Amabile (1982) that there is not a strict rule about it. The numbers may vary in accordance with the aim of the research in order to contribute to the statistical measurements, especially the inter-rater reliability. As it is cited in Cseh and Jeffries (2019) Kaufman, Plucker & Baer (2008) suggest that as it is hard to get a satisfactory level of inter-rater reliability with a small number of judges, 5-10 judges are appropriate for most of the studies. As no strict rule is presumed in guidelines, the number of the judges varies in the studies. The number in the validation studies by Amabile herself in 1982 ranges from 3 to 21. In the present study, five English teachers judged the products independently and the inter-rater reliability is calculated as high.

The CAT also provided valuable information about the creativity in the products. After the inter-rater reliability is achieved, the average of the scores gives an idea about the creativity of the products (Said-Metwaly et al., 2017). The statistical analysis shows that the mean score of the raters was 13.37 in the post-test while it was 12.11 in the pre-test. It is obvious that there is an increase after the instruction process with an implementation of

Google Translate. However, this difference is not statistically significant with a p value of .19. It is important to note that the raters judged the products of pre- and post-tests separately. Namely, those in pre-test were ranked relative to each other and the ones in post-test were ranked relative to each other. The most creative work in the pre-test might get lower points if it had been ranked relative to those in the post-test. In spite of this, there is still a rise in the post-test mean scores.

What is more, the average scores of every writing can give information about how creative the product is. While the most creative work belongs to the first student in the pre-test, the tenth student's product in the post-test is entitled as the most creative. The student who can come a long way about the creativity is the tenth student. Although he got 3.3 points in the pre-test and he got 5.6 points with a 2.3 point increase. That is why he was asked to participate in the interview. However, there are some students with slight or no changes between their tests. For example, the products of the twenty-first student got the same points, 1.6 and the thirty-third student with a 0.2-point increase.

5.2.2. Components of creativity in pre and post-tests. Although creativity involves some basic cognitive processes of thought which results in creative productions, such as divergent thinking, defining a problem or associative thinking, some standardised creativity tests, such as Torrance Test of Creative Thinking, refer to divergent thinking only (Barbot et al., 2011). The reason for this lies under the idea of Guilford that divergent thinking is a must for creativity (Guilford, 1970). As for the assessment of divergent thinking, the creative products of individuals are examined in terms of fluency, flexibility, originality and elaboration.

Thus - apart from the CAT - the products were analysed by the researcher regarding the four components of divergent thinking, namely creativity, which are identified in

Guilford's studies as fluency, flexibility, originality and elaboration. In order not to cause to statistical biases, the results of each component are discussed separately (see Section 3.9.1).

The number of the ideas in a product corresponds to the fluency score; thus every sentence in each product, regardless of being perfect in grammar or lexis, was simply counted in order to get their fluency scores. Every ideational category generated in both tests was counted to get a score for each student. It is argued by Acar, Alabbasi, Runco and Beketayev that "ideas tend to become more original and are more likely to be drawn from new conceptual categories" (Acar, Alabbasi, Runco & Beketayev, 2019, p. 2). In other words, the more ideational categories generated the more likely to be original. The originality scores corresponded to the number of unique ideas, which were not mentioned by anyone else in the group. Finally, every elaborative element in the writing was counted, such as title, conjunctions, sequencing words and some exclamation expressions to get the elaboration scores of the students.

As the homogeneity of fluency scores was not normally distributed, the median scores give an idea about the impact of using Google Translate. While it was 10 in the pre-test, it was calculated as 15 in the post-test with a 5-point increase. This difference is statistically significant with $p < .001$. It is obvious that implementing Google Translate in writing activities assisted the students to create more ideas and to express them in their second language. They were low-proficient learners and the impact of Google Translate in the fluency scores was enormous. It can be concluded that using a machine translator in writing activities as a MALL assistant in pre-editing is an effective way in flourishing the fluency of low-proficient learners.

The homogeneity of flexibility scores was calculated to be normally distributed, the mean scores were considered as an indicator. The mean score in the pre-test was $M = 5.88$ ($SD = 1.95$) and with almost a 4-point increase it was $M = 8.82$ ($SD = 1.79$) in the post-test.

The difference between these results were found to be statistically significant with a p value < .001.

The results indicate that there is an absolute increase in the number of the ideational categories of the students when they pre-edit their writing with Google Translate.

There is a great seven-point of increase in the originality of the products between the two tests, which is statistically significant. This tremendous increase is a spark in the creativity of the products. The result indicates that the low-proficient learners can generate more original ideas in writing with the help of Google Translate as a pre-edition assistant.

As for elaboration, the increase in the post-test with 2 points yields a statistically significant difference. This shows that participants can elaborate their ideas if they get help.

To sum up, Google Translate helped the participants to produce more creative work in terms of fluency, flexibility, originality and elaboration and so the first question of the present study finds an answer. Implementing machine translation in writing activities in English affects the participants' creativity in their written products positively in terms of fluency, flexibility, originality and elaboration.

These results of the study are consistent with some other studies in the field which support the idea that creativity can be sparked with the help of an appropriate training (Fontenot, N. A., 1993; Gendrop, S. C, 1996; Wang, C. W., & Horng, R. Y., 2002; Vincent, Decker, & Mumford, 2002; Scott, Leritz & Mumford, 2004; Simms, S., 2009; Vally, Z. et al, 2019).

As a means of blended learning, this study combines technology with language education and tries to find out the impact of technology on creativity. As for using a technological device in order to flourish creativity, there is a fruitful amount of studies, the results of which are in line with the present study. For example, Demiröz (2019) discusses the impact of integrating literature with technology in EFL classes on the creativity of the

students and suggests that implementing v-log, blogs, Twitter, infographics and dictation tools in literature classes enhance the creativity of EFL learners. Robin (2008) states that digital storytelling is a powerful technology in order the students to become creative story tellers. Apart from these studies, Abugohar, M. A., Yunus, K., Rabab'ah, G., & Ahmed, T. A. E. (2019) suggest that the integration of such handheld technological devices as I-pads, tablets and smart phones is a way of fostering students' creative thinking abilities. What is more, the results of a study by Anggereini, E., Budiarti, R. S., & Sanjaya, M. E. (2018) about the effect of technological tools on the students' motivation in being creative reveal that the students are more motivated to being creative when they use technological devices.

It should be highlighted that among the aforementioned studies and many others, it is difficult to find machine translators accompanied with their impact on creativity. In the literature, a tremendous amount of studies has been interested in the correlation between technology and the creativity, but not the machine translators. This present study is significant in that it fills this gap and starts a spark for further studies.

5.3. Grammatical and Lexical Knowledge

The third question of the study was about the impact of Google Translate on the participants' grammatical and lexical knowledge. In order to seek an answer for this question, the Surface Structure Taxonomy was utilised, which was first generated by Dulay et al. (1982). The taxonomy categorises the errors of language learners into four groups: omission, addition, misformation and misordering. Therefore, the products in both of the tests were analysed and the errors were detected, labelled and counted to get a score for the statistical analysis.

The differences between the pre and post-test in terms of these four error types were not statistically significant. The median scores of each test were even almost the same. It can be inferred from these findings that the students made almost the same amount of errors

although they created more amount of original sentences with more ideational categories and more details. As the creativity in their products was enhanced they wrote more grammatically accurate sentences. Although there is not a statistically significant difference between the amounts of errors, it is obvious that while their creativity was flourished, utilizing Google Translate avoided them from making errors. Hence, it can be concluded that Google Translate is an effective pre-edition tool in decreasing the grammatical and lexical knowledge of the participants in writing activities.

Taking this result into consideration, it can be said that the study is consistent with other studies about Google Translate. Lee (2019) conducted a study about the impact of using machine translation on EFL learners' writing considering the grammatical errors in the participants' writings. The participants first wrote in their first language, translated them into English on their own and then post-edit their translation by using machine translation. The two versions of translations were analysed and the study reveals that the second versions were much far beyond the first versions in that they had less lexico-grammatical errors with a statistically significant difference.

Another study on the same issue is held by Tsai (2019) with Chinese participants who were EFL learners. The procedure was a bit different from the previously mentioned study. The participants first wrote in Chinese and then drafted it in English. In the second step, they translated their Chinese text into English with the help of Google Translate and made a comparison between the first draft and the machine-translated text. Although the procedure was different the results were much or less the same. There was a progress in the machine-translated texts in terms of grammar, word-choice and spelling.

In line with the previously mentioned studies, Garcia and Pena (2011) conducted a study about Taradukka, a machine translator, and suggested that using a machine translator as a learning tool helped the participants communicate better in writing activities.

For some researchers, using machine translation is cheating and the result does not belong to the writer (Correa, 2014). However, it should be emphasised that the participants in the study wrote in their foreign language with their own effort first and they utilised Google Translate to revise their writing. That is why the products do belong to the participants. From another point of view, it is essential to mention that low-proficient learners need much help in writing as they are challenged with their lack of grammatical and lexical knowledge (Lee, 2019). Taking the results of the present study and this reality into consideration, it can be concluded that an appropriate way of utilizing Google Translate as a learning tool in EFL classes brings many opportunities ranging from flourishing creativity and reducing lexico-grammatical errors.

5.4. The Perceptions of Participants

Among the aims of the study is to find out the perceptions of high school EFL students on the use of machine translation in writing activities in English classes, in line with the third research question. In order to answer this question, six of the participants attended semi-structured interviews voluntarily and the recordings were interpreted and coded by the researcher (See Tables 12, 13, 14, 15 and 16).

It is obvious that the students are in favour of utilising Google Translate in their writing classes as all of them stated that they wanted to use it in their further language learning experiences although they were aware of the fact that it might give inaccurate output. That finding is consistent with many other studies on the use of machine translation (Alhaisoni & Alhaysony, 2017; Bahri & Mahadi, 2016; Briggs, 2018; Chandra & Yuyun, 2018; Fernández Guerra, 2014; Jolley & Maimone, 2016; Kumar, 2012; Lee, 2019; Selcuk, Jones & Vonkova, 2019; Tsai, 2019).

Another aspect of Google Translate that was mentioned by all of the interviewees is that it might give inaccurate output. It is beneficial that they are aware of the errors either

grammatical or lexical. It means that using Google Translate and facing errors increase the grammatical and lexical awareness of the learners. Editing with machine translators is resembled with peer-edition by Lee (2019, p. 12) “neither is perfect, but both are helpful to student writing”. Both ways of editing provide insight into the grammar of the target language. This result is in line with some other studies in the field (Lee, 2019; Garcia and Pena, 2011; Tsai, 2019).

Another advantage mentioned by two of the students is that they felt encouraged to write more in English. Deva mentioned that she wanted to write a book by means of Google Translate. It is promising for students to feel that enthusiasm about learning and producing, apparently they feel more motivated when they can cope with their lack of knowledge. Although there is not much research revealing the relationship between machine translators and motivation, it is evident that MALL enhances motivation in language learners (Chan, Mayall, York & Smith, 2019; Wu & Marek, 2019; Zoi, Yan & Li, 2020).

As it was mentioned before, most of the interviewees were of the opinion that using a machine translator sparked their creativity in writing. Nevertheless, one of them thought that it had nothing to do with his imagination. It is essential to express that the study did not search the ways to enhance their creative thinking ability, instead the study investigated a way to help the beginner learners produce more creative writings in their target language. The difference in between lies in the adopted creativity definition in the study which is put into words by Amabile as “The production of a novel and appropriate response, product, or solution to an open-ended task” (Amabile, 2012, p. 3). Therefore, as stated by most of the interviewees, Google Translate helped them express their novel and appropriate ideas in their writings. Although there is not much research on the impact of machine translation on creativity, the fact that beginner level learners might get efficient help from a machine

translation is evident in a study by Garcia & Pena (2011). They stated that the low proficient learners could communicate more in their target language with the help of machine translate.

As a conclusion, although the participants were aware of the fact that Google Translate might sometimes give ill-formed translations, they admitted that it provided a priceless help for them to express their original ideas in English.

5.5. Summary

As a result, the results and the discussion chapters show that all of the research questions were given an answer in a satisfactory way.

First, the analysis of the quantitative data shows that implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, originality and elaboration in a positive way. The CAT analysis shows that the creativity in the post-test products are higher although the difference is not statistically significant. What is more, the analysis of the creativity components show that there is a tremendous increase in their fluency, flexibility, originality and elaboration scores in post-test.

Second, the fact that there is not a significant difference between the pre and post-test products in terms of grammar and lexis errors reveals that the participants could produce more creative products with less errors, which answers the second question. Implementing GT in writing activities in English affects the participants' grammatical-lexical knowledge positively.

Finally, the perceptions of the students on the use of machine translation in writing activities are positive. They feel confident when they use GT in their writing and they are glad about expressing their ideas with the help GT. They feel like they learn grammatical and lexical rules efficiently.

Chapter 6

Conclusion

The ratio underlying this study is the fact that the students utilize machine translation during their foreign language learning experiences whether it is frown upon by the teachers or not. Accepting this reality, the research questions of the study emerged in pursuit of an effective way to use machine translation in writing activities. Its impact on creativity and lexico-grammatical errors is under the spotlight of the study. The questions are as follows:

RQ1. How does implementing machine translation in writing activities in English affect the participants' creativity in terms of fluency, flexibility, elaboration and originality?

RQ2. How does implementing machine translation in writing activities in English affect the participants' grammatical and lexical knowledge?

RQ3. What are the perceptions of high school EFL students on the use of machine translation in writing activities in English classes?

As it is obvious from the research questions, the study includes both quantitative and qualitative data as a mixed-method research. The data from the analysis of the written products in pre and post-tests in terms of creativity and lexico-grammatical errors form the quantitative data set and the data gathered from the semi-structured interviews about the perceptions of the participants form the qualitative data set.

The study is significant in two aspects. Firstly, there are few studies which implement a two-fold creativity assessment as the present study does. The products were analysed first by using the CAT by five raters to get an idea about the overall creativity of the products, then they were broken down into pieces in order to find the fluency, flexibility, originality and elaboration scores.

Secondly, it is unique as a study which combines creativity and machine translation. Although there are a number of studies and a huge literature about creativity and machine

translation separately, it is not common to put them together and search for the impact of machine translation on creativity in the products.

As a research with a one-group pre-post-test design, the present study reveals some valuable information about the impact of machine translation on the creativity of the low-proficient EFL learners in writing. The CAT revealed a high inter-rater reliability, which is the most essential part of this technique and a rank-order of the products became possible thanks to this technique. Additionally, the analysis about the components of creativity revealed that the use of machine translation (Google Translate in this study) as a MALL tool in pre-editing process in writing enhances the creativity in the products of low-proficient learners, who are limited with a small amount of grammatical and lexical knowledge. The increase in the scores of the four components of creativity – fluency, flexibility, originality and elaboration – was statistically significant.

It should definitely be put into words that the grammatical and lexical errors of the participants did not increase in the post-test although it was proved that they produced more amount of sentences with original ideas. Though there was not a statistically significant difference between the scores of the two tests, it is obvious that the participants produced more accurate sentences by means of Google Translate. It is proven that when Google Translate is applied as a pre-edition assistant in classroom activities, the learners can produce more amount of sentences with less errors.

Finally, the semi-structured interviews revealed that the participants were in favour of utilising Google Translate even it sometimes translated ill-formed sentences. They mentioned that dealing with those errors raised their grammatical and lexical awareness in their target language. In addition, they felt more encouraged to create new written products with the help of Google Translate as they were aware that the creativity in the post-tests products was higher than the pre-test products. All of them reported that they would continue utilising

Google Translate as it was beneficial in their language learning process as a free and accessible tool.

6.1. Implications of the Study for EFL Teachers

As stated by Prensky (2001) our students are digital natives who are born into technology and they spend more time with technological devices than they do with pens and papers. Thus, he suggests the teachers who are digital immigrants to change the way they teach and to keep up with the students' needs. Bearing this in mind, we, as teachers, cannot turn a blind eye on technological tools which open new horizons for EFL learners by giving them the opportunity to be more creative with the help of multimodal items such as sounds, visuals, animations or graphics (Yoon, 2013).

The upcoming requirements of digital societies force the individuals to be more creative to keep up with the rapid developments in every field of our life. The field of education has inevitably been affected by these dramatic changes. Beforehand bringing a tape recorder to the class might be enough to attract students' attention, today it is impossible to keep them alert enough to have them participate in the classroom activities.

Technology poses as one of the key issues in meeting needs of new generation who simply reject and make fun of the existing education system, which was formed in the era of industrial revolution. However, as far as the traditional education system, which provides an image of a classroom with desks in line and students sitting passively, is insisted not to be revolved in terms of the digital era's needs, our students may not be able to gain the required abilities for a revolutionary future, let alone being creative.

The future calls for new technologies and new digital abilities. Our students can keep up with these rapid changes only if we enlarge the scope of our lessons by embracing the merits of every technological tool. As this present research proves Google Translate can serve as an effective learning tool in raising creativity in low-proficient learners' written products

and in decreasing the number of grammatical and lexical errors when it is used as a MALL assistant in pre-editing.

What is more, it is inferred from the literature that students are in favour of using machine translate in their language learning experience no matter we ban or not. Then, we should show them how they can use it effectively and make benefit out of it. As the present study suggests writing activities can be arranged with Google Translate in order to help our students create their own products with less errors. In this way, the effective use of Google Translate also sheds light on them about being aware of grammatical rules. Knowing that they feel more encouraged to write in English with Google Translate as a pre-editing tool assistant is worth every effort in teaching them.

Finally, those who want to integrate Google Translate into EFL classes should be aware of the fact that GT is designed as a machine translator, not a learning tool. It is quite possible to get ill-formed translations. Thus, they should design the objectives and the procedures in their classes carefully enough in order not to result in a fossilisation of errors due to the mistranslated outputs from GT. The students also should be given a training session about such possible errors and how to avoid them with simple regulations like using initials and punctuation marks accurately.

6.2. Suggestions for Further Research

The present study is one of the unique studies the participants of which is low-proficient language learners. There is a need for further research about this group of participants and their perspectives on the implementation of machine translation as a MALL tool in their classrooms.

Furthermore, as the students are not allowed to bring their mobile phones to the classrooms in state schools in accordance with the related regulations, the perceptions of the parents, the school principals and the legislators on MALL should be investigated. The

researchers should get a dispensation to administer research on MALL, which poses possible unexpected risks during the instruction process. It is urgent to reveal scientific information about the impact of every technological tool in order to regulate such rules.

In addition, the fact that there is little research on machine translation and its impact on creativity calls for new research. This promising gap in the junction point of these two huge subjects needs more attention from different aspects: teachers' or learners' attitudes, their perception, other effective ways of using machine translation in writing, creativity in students' products with different levels.

References

Abedi , J. (2000). *Abedi test of creativity (ATC)*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles.

Abugohar, M. A., Yunus, K., Rabab'ah, G., & Ahmed, T. A. E. (2019). Integrating cloud world synergy in ELT to adults: Perspectives of handheld technologies. *International Journal of Interactive Mobile Technologies (iJIM)*, *13*(10), 150-168. https://doi.org/10.3991/ijim.v13i10.10783

Acar, S., Alabbasi, A. M. A., Runco, M. A., & Beketayev, K. (2019). Latency as a predictor of originality in divergent thinking. *Thinking Skills and Creativity*, *33*. https://doi.org/10.1016/j.tsc.2019.100574

Albert, A., & Kormos, J. (2011). Creativity and narrative task performance: An exploratory study. *Language Learning*, *61*, 73-99. https://doi.org/10.1111/j.1467-9922.2011.00643.x

Alhaisoni, E., & Alhaysony, M. (2017). An investigation of Saudi EFL university students' attitudes towards the use of Google Translate. *International Journal of English Language Education*, *5*(1), 72-82. doi:10.5296/ijele.v5i1.10696

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, *43*(5), 997. https://doi.org/10.1037/0022-3514.43.5.997

Amabile, T. M. (1983). The social psychology of creativity: A componential conceptualization. *Journal of Personality and Social Psychology*, *45*(2), 357. https://doi.org/10.1037/0022-3514.45.2.357

Amabile, T. M. (1988). A model of creativity and innovation in organizations. *Research in Organizational Behavior*, *10*(1), 123-167. Retrieved from http://web.mit.edu

Amabile, T. M. (2012). Componential theory of creativity. *Harvard Business School*, *12*(96), 1-10. Retreived from https://pdfs.semanticscholar.org/6188/5f52d813d518b4ed5b833b4022990211f063.pdf

Amabile, T. M. (2018). *Creativity in context: Update to the social psychology of creativity*. Routledge. https://doi.org/10.4324/9780429501234

Anggereini, E., Budiarti, R. S., & Sanjaya, M. E. (2018). The effect of ICT-based classwide peer tutoring (CWPT) application and motivation to biology education student creativity in ICT learning. *BIODIK*, *4*(2), 105-113. https://doi.org/10.22437/bio.v4i2.6177

Baer, J. (1994). Performance assessments of creativity: Do they have long-term stability?. *Roeper Review*, *17*(1), 7-11. https://doi.org/10.1080/02783199409553609

Baer, J. (2003). The impact of the core knowledge curriculum on creativity. *Creativity Research Journal*, *15*(2-3), 297-300. https://doi.org/10.1080/10400419.2003.9651422

Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the consensual assessment technique to nonparallel creative products. *Creativity Research Journal*, *16*(1), 113-117. https://doi.org/10.1207/s15326934crj1601_11

Baer, J., & McKool, S. S. (2009). Assessing creativity using the consensual assessment technique. In C. S. Schreiner (Ed.), *Handbook of research on assessment technologies, methods, and applications in higher education* (pp. 65–77). PA: Hershey. doi: 10.4018/978-1-60566-667-9.ch004

Bahri, H., & Mahadi, T. S. T. (2016). The application of mobile devices in the translation classroom. *Advances in Language and Literary Studies*, *7*(6), 237-242. Retrieved from https://eric.ed.gov/?id=EJ1126945

Ball, O. E., & Torrance, E. P. (1984). *Streamlines scoring workbook: Figural A*. Bensenville, IL: Scholastic Testing Service Inc.

Barbot, B., Besançon, M., & I Lubart, T. (2011). Assessing creativity in the classroom. *The Open Education Journal*, *4*(1). doi: 10.2174/1874920801104010058

Barkaoui, K., & Knouzi, I. (2012). Combining score and text analyses to examine task equivalence in writing assessment. In E. Sttendam, M. Tillerma, G. Rijlaarsdam, & H. Bergh (Eds.). *Measuring writing: Recent insights into theory, methodology and practices* (pp. 83–116). Boston: Davers.

Bartholomew, S. R., & Reeve, E. (2018). Middle school student perceptions and actual use of mobile devices: Highlighting disconnects in student planned and actual usage of mobile devices in class. *Journal of Educational Technology & Society, 21*(1), 48-58. Available on https://www.jstor.org/stable/26273867.

Beghetto , R. A. (2010). Creativity in the classroom. In J. C. Kaufman & R. J. Sternberg (Eds.). *Cambridge handbook of creativity* (pp. 447 – 463). New York, NY: Cambridge University Press.

Behr, M. J. (1970). Interactions between "Structure-of-Intellect" factors and two methods of presenting concepts of modular arithmetic: A summary paper. *Journal for Research in Mathematics Education*, 29-42.

Boden, M. 2001. Creativity & knowledge. In A. Craft, B. Jeffery and M. Leibling (Eds.). *Creativity in education* (pp. 95 – 102). London: Continuum.
Retrieved from https://books.google.com.tr/

Briggs, N. (2018). Neural machine translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal*, *14*(1), 2-24.

Btoosh, M. A. (2011). Error types in the approximate system of Arab students of English: A multiple classificatory taxonomy. *Indian Journal of Applied Linguistics*, *37*(1).

Bulut, E., Baydar Ertopcu, F. Umut Ozardali, S. & Senturk, S. (2019). *Secondary Education Teenwise 9: Student's book*. Ankara: Promat Basım Yayın.

Calis, E., & Dikilitas, K. (2012). The use of translation in EFL classes as L2 learning practice. *Procedia-Social and Behavioral Sciences*, *46*, 5079-5084. https://doi.org/10.1016/j.sbspro.2012.06.389

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books. Available on http://www.sfu.ca/~palys/Campbell&Stanley-1959-Exptl&QuasiExptlDesignsForResearch.pdf

Caruth, G. D. (2013). Demystifying mixed methods research design: A review of the literature. *Mevlana International Journal of Education*, *3*(2), 112-122. http://dx.doi.org/10.13054/mije.13.35.3.2

Case, M. (2015). Machine translation and the disruption of foreign language learning activities. *eLearning Papers*, (45), 4-16. Available on http://www.divaportal.org/smash/record.jsf?pid=diva2%3A874792&dswid=-2180

Chandra, S. O., & Yuyun, I. (2018). The use of google translate in EFL essay writing. *LLT Journal: A Journal on Language and Language Teaching*, *21*(2), 228-238. doi: doi.org/10.24071/llt.2018.210212

Chen, Y., Mayall, H. J., York, C. S., & Smith, T. J. (2019). Parental perception and English learners' mobile-assisted language learning: An ethnographic case study from a technology-based Funds of Knowledge approach. *Learning, Culture and Social Interaction*, *22*, 100325. https://doi.org/10.1016/j.lcsi.2019.100325

Chomsky, N. (1968). Language and mind. *Psychology Today*, *1*(9), 48. Available on https://books.google.com.tr/

Clifford, J., Merschel, L., & Munné, J. (2013). Surveying the landscape: What is the role of machine translation in language learning? *Research in Education and Learning Innovation Archives, 10*, 108-121. https://doi.org/10.7203/attic.10.2228

Conti, R., Coon, H., & Amabile, T. M. (1996). Evidence to support the componential model of creativity: Secondary analyses of three studies. *Creativity Research Journal*, *9*(4), 385-389. https://doi.org/10.1207/s15326934crj0904_9

Correa, M. (2014). Leaving the "peer" out of peer-editing: Online translators as a pedagogical tool in the Spanish as a second language classroom. *Latin American Journal of Content & Language Integrated Learning*, *7*(1), 1-20. doi: https://doi.org/10.5294/3568

Costa, Â., Ling, W., Luís, T., Correia, R., & Coheur, L. (2015). A linguistically motivated taxonomy for Machine Translation error analysis. *Machine Translation*, *29*(2), 127-161. doi: 10.1007/s10590-015-9169-0

Craft, A. (2001). An analysis of research and literature on creativity in education. *Qualifications and Curriculum Authority*, *51*(2), 1-37. Available on https://pdfs.semanticscholar.org/b26f/f91d08b36522a301e23c84130e8922284e26.pdf

Cramond, B., Matthews-Morgan, J., Bandalos, D., & Zuo, L. (2005). A report on the 40-year follow-up of the Torrance Tests of Creative Thinking: Alive and well in the new millennium. *Gifted Child Quarterly*, *49*(4), 283-291. https://doi.org/10.1177/001698620504900402

Creativity. (n. d.). In *Cambridge online dictionary*. Retrieved from https://dictionary.cambridge.org/tr/sözlük/ingilizce-türkçe/creative?q=creativity

Creativity. (n. d.). In *Longman dictionary of contemporary English online*. Retrieved from https://www.ldoceonline.com/dictionary/creativity

Creativity. (n. d.). In *Oxford online learner's dictionary*. Retrieved from https://www.oxfordlearnersdictionaries.com/definition/english/creativity?q=creativity

Cseh, G. M., & Jeffries, K. K. (2019). A scattered CAT: A critical evaluation of the consensual assessment technique for creativity research. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(2), 159. http://dx.doi.org/10.1037/aca0000220

Demiröz, H. (2019). Yaratıcılığı arttırmak için teknoloji ve edebiyatin dil öğretiminde birleştirilmesi. *Ana Dili Eğitimi Dergisi*, *7*(1), 65-78. https://doi.org/10.16916/aded.448839

Dikici, A. (2013). The adaptation of creativity fostering primary teachers index scale into Turkish. *Educational Sciences: Theory and Practice, 13*(1), 318-323.

Dörnyei, Z. (2011). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Drapeau, P. (2014). *Sparking student creativity: Practical ways to promote innovative thinking and problem solving*. Alexandria, VA: Association for Supervision and Curriculum Development. Retrieved from https://books.google.com.tr.

Ducar, C., & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical solutions for making peace with Google Translate. *Foreign Language Annals, 51*(4), 779-795. doi: 10.1111/flan.12366

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford: Oxford University Press, Retrieved from https://eric.ed.gov/?id=ED227699

Ekinci, E., & Ekinci, M. (2017). Perceptions of EFL learners about using mobile applications for English language learning: A case study. *International Journal of Language Academy*, *5*(5), 175-193. http://dx.doi.org/10.18033/ijla.3659

Feldman, D. H., Csikszentmihalyi, M., & Gardner, H. (1994). *Changing the world: A framework for the study of creativity.* Westport, CT: Praeger.

Fernández Guerra, A. B. (2014). The usefulness of translation in foreign language learning: Students' attitudes. *International Journal of English Language & Translation Studies, 2*(1), 153-170. Retrieved from http://www.eltsjournal.org

Fontenot, N. A. (1993). Effects of training in creativity and creative problem finding upon business people. *The Journal of Social Psychology*, *133*(1), 11-22. https://doi.org/10.1080/00224545.1993.9712114

Fromkin, V., Rodman, R., & Hyams, N. (2018). *An introduction to language*. Boston, MA: Cengage Learning. Available on http://relin.letras.ufmg.br/shlee/Fromkinch1.pdf

Garcia, I., & Pena, M. I. (2011). Machine translation-assisted language learning: writing for beginners. *Computer Assisted Language Learning*, *24*(5), 471-487. doi: 10.1080/09588221.2011.582687

Garrido, C. G., & Rosado Romero, C. (2012). Errors in the use of English tenses. *Íkala, revista de lenguaje y cultura*, *17*(3), 285-296. Available on http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0123-34322012000300006

Gendrop, S. C. (1996). Effect of an intervention in synectics on the creative thinking of nurses. *Creativity Research Journal, 9*(1), 11-19. https://doi.org/10.1207/s15326934crj0901_2

Goff, K. (2002). *Abbreviated Torrance test for adults: Manual*. Bensenville, IL: Scholastic Testing Service.

Gonçalves, F., Cabral, D., Campos, P., & Schöning, J. (2017). I smell creativity: exploring the effects of olfactory and auditory cues to support creative writing tasks. In *IFIP Conference on Human-Computer Interaction* (pp. 165-183). Springer, Cham. https://doi.org/10.1007/978-3-319-67684-5_11

Goodlad, J. L. (2004). *A place called school: Prospects for the future*. New York: McGrawHill.

Grainger, T., Goouch, K., & Lambirth, A. (2005). *Creativity and writing: Developing voice and verve in the classroom*. London: Routledge.

Gralewski, J., & Karwowski, M. (2018). Are teachers' implicit theories of creativity related to the recognition of their students' creativity?. *The Journal of Creative Behavior*, *52*(2), 156-167. https://doi.org/10.1002/jocb.140

Gribbons, B., & Herman, J. (1996). True and quasi-experimental designs. *Practical Assessment, Research, and Evaluation*, *5*(1), 14. Available on https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1069&context=pare

Groves, M., & Mundt, K. (2015). Friend or foe? Google Translate in language for academic purposes. *English for Specific Purposes*, *37*, 112-121. http://dx.doi.org/10.1016/j.esp.2014.09.001

Guilford, J. P. (1950). *Creativity*. American psychology.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin*, *53*(4), 267. https://doi.org/10.1037/h0040755

Guilford, J. P. (1966). Measurement and creativity. *Theory into Practice*, *5*(4), 185-189. https://doi.org/10.1080/00405846609542023

Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14. https://doi.org/10.1002/j.2162-6057.1967.tb00002.x

Guilford, J. P. (1970). Creativity: Retrospect and prospect. *The Journal of Creative Behavior*, *4*(3), 149-168. https://doi.org/10.1002/j.2162-6057.1970.tb00856.x

Gursoy, E., & Bag, H. K. (2018). Is it possible to enhance the creative thinking skills of EFL learners through training?. *Advances in Language and Literary Studies*, *9*(6), 172-182. http://dx.doi.org/10.7575/aiac.alls.v.9n.6p.172

Hennessey, B. A. (1994). The consensual assessment technique: An examination of the relationship between ratings of product and process creativity. *Creativity Research Journal*, *7*(2), 193-208. https://doi.org/10.1080/10400419409534524

Hennessey, B. A., Amabile, T. M., & Mueller, J. S. (1999). Consensual assessment. *Encyclopedia of Creativity*, *1*, 346-359. Available on https://books.google.com.tr/

Henriksen, D., Mishra, P., & Fisser, P. (2016). Infusing creativity and technology in 21st century education: A systemic view for change. *Journal of Educational Technology & Society*, *19*(3), 27-37.

Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, *49*(3), 234-244. https://doi.org/10.2307/3345709

Hocevar, D. (1981). Measurement of creativity: Review and critique. *Journal of Personality Assessment*, *45*(5), 450-464. https://doi.org/10.1207/s15327752jpa4505_1

Hocevar, D., & Michael, W. (1979). The effects of scoring formulas on the discriminant validity of tests of divergent thinking. *Educational and Psychological Measurement, 39*, 917–921. https://doi.org/10.1177/001316447903900427

Hockly, N. (2018). Blended learning. *ELT Journal*, *72*(1), 97-101. https://doi.org/10.1093/elt/ccx058

Irfan, H. (2018). *Testing creative writing in Pakistan*. Cambridge Scholars Publishing. Retrieved from https://books.google.com.tr

Jin, L., & Deifell, E. (2013). Foreign language learners' use and perception of online dictionaries: A survey study. *Journal of Online Learning and Teaching*, *9*(4), 515. Retrieved from https://jolt.merlot.org/vol9no4/jin_1213.pdf

Jolley, J. R., & Maimone, L. (2015). Free online machine translation: Use and perceptions by Spanish students and instructors. In A. J. Moeller (Ed.). *Learn languages, explore cultures, transform lives* (pp. 181-200). Minneapolis: 2015 Central States Conference on the Teaching of Foreign Languages. Available on https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1298&context=teachlearnfacpub#page=202

Karwowski, M. (2007). Teachers' nominations of students' creativity: Should we believe them? Are the nominations valid?. *Social Sciences, 2,* 264-269. Available on http://docsdrive.com/pdfs/medwelljournals/sscience/2007/264-269.pdf

Karwowski, M., Gralewski, J., & Szumski, G. (2015). Teachers' effect on students' creative self-beliefs is moderated by students' gender. *Learning and Individual Differences, 44,* 1-8. https://doi.org/10.1016/j.lindif.2015.10.001

Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment*. New York: Wiley. Available on https://books.google.com.tr

Kaygın, B., & Çetinkaya, Ç. (2015). Yaratıcılığın değerlendirmesinde yeni yaklaşımlar. *Üstün Zekalılar Eğitimi ve Yaratıcılık Dergisi, 2*(1), 1-11. Retrieved from https://dergipark.org.tr/tr/pub/jgedc/issue/38679/449349

Kırkgöz, Y. (2014). Exploring poems to promote language learners' creative writing. *Procedia-Social and Behavioral Sciences, 158*, 394-401. https://doi.org/10.1016/j.sbspro.2014.12.106

Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal, 18*(1), 3-14. https://doi.org/10.1207/s15326934crj1801_2

Kliffer, M. (2005). An experiment in MT post-editing by a class of intermediate/advanced French majors. In *Proceedings EAMT 10th Annual Conference* (pp. 160-165).

Kolokoltsev, M., Iermakov, S., & Potop, V. (2017). Creativity project method in academic physical education process. *Journal of Physical Education and Sport*, *17*(4), 2659-2664. doi:10.7752/jpes.2017.04305

Kozbelt, A., Beghetto, R. A., & Runco, M. A. (2010). Theories of creativity. In Kaufman, J. C. & Sternberg, R. J. (Eds.), *The Cambridge handbook of creativity* (pp. 20 – 47). Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511763205.004

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Kumar, A. (2012). Machine translation in Arabic-speaking ELT classrooms: Applications and implications. *International Journal of Social Science and Humanity*, *2*(6), 442-445. Retrieved from https://www.researchgate.net/profile/Avinash_Kumar114/publication/269939422_Machine_Translation_in_Arabic_Speaking_ELT_Classrooms_Applications_and_Implications/links/5728d26b08aef7c7e2c0c849/Machine-Translation-in-Arabic-Speaking-ELT-Classrooms-Applications-and-Implications.pdf

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics,* 159-174. doi: 10.2307/2529310

Lee, S. M. (2019). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 1-19. https://doi.org/10.1080/09588221.2018.1553186

Leonardi, V. (2010). *The role of pedagogical translation in second language acquisition: From theory to practice*. Bern: Peter Lang. Retrieved from https://books.google.com.tr/

Maley, A., & Bolitho, R. (2015). Creativity. *ELT Journal*, *69*(4), 434-436. https://doi.org/10.1093/elt/ccv036

Maley, A., & Peachy, N. (Eds.). (2015). *Creativity in the English language classroom*. London: British Council.

Maniam, M., & Rajagopal, P. (2016). Simple past tense errors based on surface structure taxonomy in ESL Malaysian undergraduates writing. *Global Journal of Advanced Research*, *3*(6), 547-553.

McDonough, K., Crawford, W. J., & Mackey, A. (2015). Creativity and EFL students' language use during a group problem-solving task. *TESOL Quarterly*, *49*(1), 188-199.

Ministry of National Education. (2018). 9^{th} – 12^{th} grades English curriculum. 1-69. Retrieved from http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=342

Moran, S. (2010). The roles of creativity in society. In Kaufman, J. C. & Sternberg, R. J. (Eds.), *Cambridge handbook of creativity* (pp. 74 – 90). New York, NY: Cambridge University Press. Available on https://books.google.com.tr.

Morse, J. (2009). *Mixed method design: Principles and procedures*. New York: Routledge, https://doi.org/10.4324/9781315424538

Naderi, H., Abdullah, R., Aizan, H. T., Sharir, J., & Kumar, V. (2009). Creativity, age and gender as predictors of academic achievement among undergraduate students. *Journal of American Science*, *5*(5), 101-112.

Niño, A. (2008). Evaluating the use of machine translation post-editing in the foreign language class. *Computer Assisted Language Learning*, *21*(1), 29-49. https://doi.org/10.1080/09588220701865482

Niño, A. (2009). Machine translation in foreign language learning: language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, *21*(2), 241-258. doi: https://doi.org/10.1017/S0958344009000172

Nusbaum, E. C., & Silvia, P. J. (2011). Are intelligence and creativity really so different?: Fluid intelligence, executive processes, and strategy use in divergent thinking. *Intelligence*, *39*(1), 36-45. https://doi.org/10.1016/j.intell.2010.11.002

Okpor, M. D. (2014). Machine translation approaches: issues and challenges. *International Journal of Computer Science Issues (IJCSI)*, *11*(5), 159.

O'Neill, E. (2012). *The effect of online translators on L2 writing in French* (Doctoral dissertation). University of Illinois at Urbana-Champaign. Retrieved from http://hdl.handle.net/2142/34317

Ozer, O., & Kilic, F. (2018). The effect of mobile-assisted language learning environment on EFL students' academic achievement, cognitive load and acceptance of mobile learning tools. *EURASIA Journal of Mathematics, Science and Technology Education, 14*(7), 2915-2928. Retrieved from https://dergipark.org.tr/tr/download/article-file/199876

Özerbaş, M. A., & Mart, Ö. A. (2017). İngilizce öğretmen adaylarının Web 2.0 kullanımına ilişkin görüş ve kullanım düzeyleri. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, *18*(3).

Phonethibsavads, A., Bender, S., & Peppler, K. (2019). Utilizing the consensual assessment technique to compare creativity in drama spaces. *Creativity. Theories–Research-Applications*, *6*(1), 4-19. doi: 10.1515/ctra-2019-0001

Pishghadam, R., Baghaei, P., & Shayesteh, S. (2012). Construction and validation of an English language teacher creativity scale (ELT-CS). *Journal of American Science, 8*(3), 497-508.

Pishghadam, R., & Mehr, F. J. (2011). Learner creativity and performance in written narrative tasks. *World Journal of Education*, *1*(2), 115-125. http://dx.doi.org/10.5430/wje.v1n2p115

Pope, R. (2005). *Creativity: Theory, history, practice*. New York: Routledge.

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon*, *9*(5), 1-6. https://doi.org/10.1108/10748120110424816

Rehm, G., Sasaki, F., Stein, D., & Witt, A. (Eds.). (2018). *Language technologies for a multilingual Europe: TC3 III*. Berlin: Language Science Press. doi:10.5281/zenodo.1291947

Robin, B. R. (2008). Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory into Practice*, *47*(3), 220-228. doi: 10.1080/00405840802153916

Rubinstein, G. (2003). Authoritarianism and its relation to creativity: A comparative study among students of design, behavioral sciences and law. *Personality and Individual Differences*, *34*(4), 695-705. https://doi.org/10.1016/S0191-8869(02)00055-7

Runco, M. A. (1986). Flexibility and originality in children's divergent thinking. *The Journal of Psychology*, *120*(4), 345-352. https://doi.org/10.1080/00223980.1986.9712632

Runco, M. A. (2004). Creativity. *Annual Review of Psychology*, 55(1), 657-687. doi: 10.1146/annurev.psych.55.090902.141502

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal*, *24*(1), 66-75. https://doi.org/10.1080/10400419.2012.652929

Runco, M. A., & Albert, R. S. (1985). The reliability and validity of ideational originality in the divergent thinking of academically gifted and non-gifted children. *Educational and Psychological Measurement, 45*(3), 483-501.

Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, *24*(1), 92-96. https://doi.org/10.1080/10400419.2012.650092

Runco, M. A. (2014). *Creativity: Theories and themes: Research, development, and practice*. Boston: Elsevier. Retrieved from https://books.google.com.tr.

Saad, M. A. H., & Sawalmeh, M. H. M. (2014). Error analysis in role-play presentations among less proficient L2 Malaysian learners. *International Journal of English Education, 3*(3), 170-181. Retrieved from http://ir.unikl.edu.my/jspui/handle/123456789/10242

Said-Metwaly, S., Van den Noortgate, W., & Kyndt, E. (2017). Approaches to measuring creativity: a systematic literature review. *Creativity. Theories–Research-Applications*, *4*(2), 238-275. doi: 10.1515/ctra-2017-0013

Sangmin-Michelle Lee (2019): The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 1-19. doi:10.1080/09588221.2018.1553186

Saroukhil, M. A., Ghalkhani, O., & Hashemi, A. (2018). A critical review of translation: A look forward. *International Journal of Education and Literacy Studies*, *6*(2), 101-110. doi: http://dx.doi.org/10.7575/aiac.ijels.v.6n.2p.101

Scott, D., & Morrison, M. (2006). *Key ideas in educational research*. London: Continuum. Available on https://books.google.com.tr

Scott, G., Leritz, L. E., & Mumford, M. D. (2004). The effectiveness of creativity training: A quantitative review. *Creativity Research Journal*, *16*(4), 361-388. https://doi.org/10.1080/10400410409534549

Selcuk, H., Jones, J. & Vonkova, H. (2019). The use of Google Translate as an ICT tool in web-based collaborative writing: Self-reported accounts of EFL learners. In H. Kratochvilova & R. Kratochvil (Eds.), *Proceedings of IAC 2019 in Vienna* (pp. 74-80). Prague: Czech Institute of Academic Education.

Shei, C. C. (2002). Teaching MT through pre-editing: Three case studies. In *6th EAMT Workshop Teaching Machine Translation* (pp. 89-98). Retrieved from http://mt-archive.info/EAMT-2002-Shei.pdf

Silvia, P. J., Martin, C., & Nusbaum, E. C. (2009). A snapshot of creativity: Evaluating a quick and simple method for assessing divergent thinking. *Thinking Skills and Creativity*, *4*(2), 79-85. https://doi.org/10.1016/j.tsc.2009.06.005

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., ... & Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts*, *2*(2), 68. https://doi.org/10.1037/1931-3896.2.2.68

Simms, S. (2009). The impact of creativity training on an accounting negotiation. *Group Decision and Negotiation*, *18*(1), 75-87. doi: 10.1007/s10726-008-9124-z

Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 3–15). New York: Cambridge University Press. Available on https://books.google.com.tr/

Sukkhwan, A. (2014). *Students' attitudes and behaviors towards the use of google translate* (Doctoral dissertation). Prince of Songkla University. Retrieved from http://kb.psu.ac.th/psukb/bitstream/2010/9459/1/387714.pdf

Tashakkori, A., & Teddlie, C. (Eds.). (2010). *Sage handbook of mixed methods in social & behavioral research*. California: Sage. Available on https://books.google.com.tr

Thompson, C. B., & Panacek, E. A. (2006). Research study designs: experimental and quasi-experimental. *Air Medical Journal*, *25*(6), 242-246. https://doi.org/10.1016/j.amj.2006.09.001

Tin, T. B. (2015). Creativity and creative tasks for language learners. *Proceedings of ISELT FBS Universitas Negeri Padang*, *3*, 43-44. Retrieved from http://ejournal.unp.ac.id/index.php/selt/article/view/6809/5346

Tomlinson, B., & Whittaker, C. (2013). *Blended Learning in English Language Teaching*. London: British Council.

Torrance, E. P. (1968). A longitudinal examination of the fourth grade slump in creativity. *Gifted Child Quarterly*, *12*(4), 195-199. https://doi.org/10.1177/001698626801200401

Torrance, E. P. (1974). *The Torrance tests of creative thinking-TTCT manual and scoring guide: Verbal test A, figural test*. Lexington, KY: Ginn.

Torrance, E. P. (1990). *The Torrance tests of creative thinking norms—technical manual figural (streamlined) forms a & b*. Bensenville, IL: Scholastic Testing Service, Inc

Torrance, E. P. (1988). The nature of creativity as manifest in its testing. *The Nature of Creativity*, 43-75. Retrieved from https://books.google.com.tr/

Torrance, E. P. (1990). *Technical supplement: Verbal forms a and b for the Torrance tests of creative thinking*. Bensonville, IL: Scholastic Testing Service.

Torrance, E. P. (1998). *The Torrance tests of creative thinking norms—technical manual figural (streamlined) forms a & a*. Bensenville, IL: Scholastic Testing Service, Inc.

Treffinger, D. J., Young, G. C., Selby, E. C., & Shepardson, C. (2002). Assessing creativity: A guide for educators. *National Research Center on the Gifted and Talented*. Retrieved from https://files.eric.ed.gov/fulltext/ED505548.pdf

Tsai, S. C. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning,* 1-17. doi: 10.1080/09588221.2018.1527361

Uysal, M. E. (2009). *İlköğretim Türkçe dersinde işbirlikli öğrenmenin erişi, eleştirel düşünce ve yaratıcılık becerilerine etkisi* (Doctoral dissertation). DEÜ Eğitim Bilimleri Enstitüsü. Retrieved from http://hdl.handle.net/20.500.12397/6824

Vally, Z., Salloum, L., AlQedra, D., El Shazly, S., Albloshi, M., Alsheraifi, S., & Alkaabi, A. (2019). Examining the effects of creativity training on creative production, creative self-efficacy, and neuro-executive functioning. *Thinking Skills and Creativity*, *31*, 70-78. https://doi.org/10.1016/j.tsc.2018.11.003

Vincent, A. S., Decker, B. P., & Mumford, M. D. (2002). Divergent thinking, intelligence, and expertise: A test of alternative models. *Creativity Research Journal*, *14*(2), 163-178. https://doi.org/10.1207/S15326934CRJ1402_4

Wallas, G. (1926). *The art of thought.* London: Jonathan Cape.

Wang, C. W., & Horng, R. Y. (2002). The effects of creative problem solving training on creativity, cognitive type and R&D performance. *R&D Management*, *32*(1), 35-45. https://doi.org/10.1111/1467-9310.00237

Wang, H. C. (2019). Fostering learner creativity in the English L2 classroom: Application of the creative problem-solving model. *Thinking Skills and Creativity*, *31*, 58-69. https://doi.org/10.1016/j.tsc.2018.11.005

Weaver, W. (1955). Translation. *Machine Translation of Languages*, *14*, 15-23. Retrieved from http://htl.linguist.univ-paris-diderot.fr/_media/biennale/et09/supportscours/leon/leon.pdf

White, K. D., & Heidrich, E. (2013). Our policies, their text: German language students' strategies with and beliefs about web-based machine translation. *Die Unterrichtspraxis/Teaching German*, *46*(2), 230-250. https://doi.org/10.1111/tger.10143

Wu, P. H. N., & Marek, M. W. (2020). Designing interactive cross-cultural mobile-assisted language learning. In Khosrow-Pour, M. (ed.), *Multicultural instructional design: Concepts, methodologies, tools, and applications* (pp. 452-475). IGI Global. doi: 10.4018/978-1-5225-9279-2.ch021

Yoon, T. (2013). Are you digitized? Ways to provide motivation for ELLs using digital storytelling. *International Journal of Research Studies in Educational Technology*, *2*(1), 1-10. https://doi.org/10.5861/ijrset.2012.204

Yurdagül, C., & Öz, S. (2018). Attitude towards mobile learning in English language education. *Education Sciences, 8*(3), 142. https://doi.org/10.3390/educsci8030142

Zou, B., Yan, X., & Li, H. (2020). Students' perspectives on using online sources and apps for EFL learning in the mobile-assisted language learning context. In Khosrow-Pour, M (eds.). *Language learning and literacy: Breakthroughs in research and practice* (pp. 515-531). IGI Global. doi: 10.4018/978-1-5225-9618-9.ch027

Appendices

Appendix 1: Background Information Form

Öğrenci Tanıma Formu

Lütfen, gönüllü olarak katıldığınız ‘‘Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi’’ başlıklı araştırma çalışması için hazırlanmış bu formu doldurunuz. Bilgileriniz tamamen gizli tutulacaktır.

• Kişisel Bilgiler

1) Cinsiyetiniz ( ) Kız ( ) Erkek

2) Doğum tarihiniz …. / …. / ……

2) Anadiliniz nedir? _____

3) Hangi ülkede doğdunuz? _____

4) Doğdunuz ülkeden başka herhangi bir ülkede bulundunuz mu? Eğer cevabınız “Evet” ise lütfen neresi olduğunu, neden ve ne kadar süre orada kaldığınızı belirtiniz.

_____

_____

4) İngilizce de dâhil olmak üzere hangi dilleri biliyorsunuz ya da şimdiye kadar öğrenmeye çalıştınız?

Dil	Ne zaman öğrendiniz?	Nerede öğrendiniz?	Ne kadar akıcı konuşabilirsiniz? (hiç/az/orta/iyi/mükemmel)

5) Sizce İngilizce öğreniminizde yaratıcılığın etkisi nedir? Lütfen birkaç cümleyle açıklayınız.

_____

_____

_____

_____

6) İngilizce öğrenme konusunda kendinizi motive hissediyor musunuz? ( ) Evet ( ) Hayır

7) İngilizceye dair gelecek planlarınız nelerdir? (Size uyduğunu düşündüklerinizi işaretleyiniz.)

____ Meslek hayatımda işime yarayacak

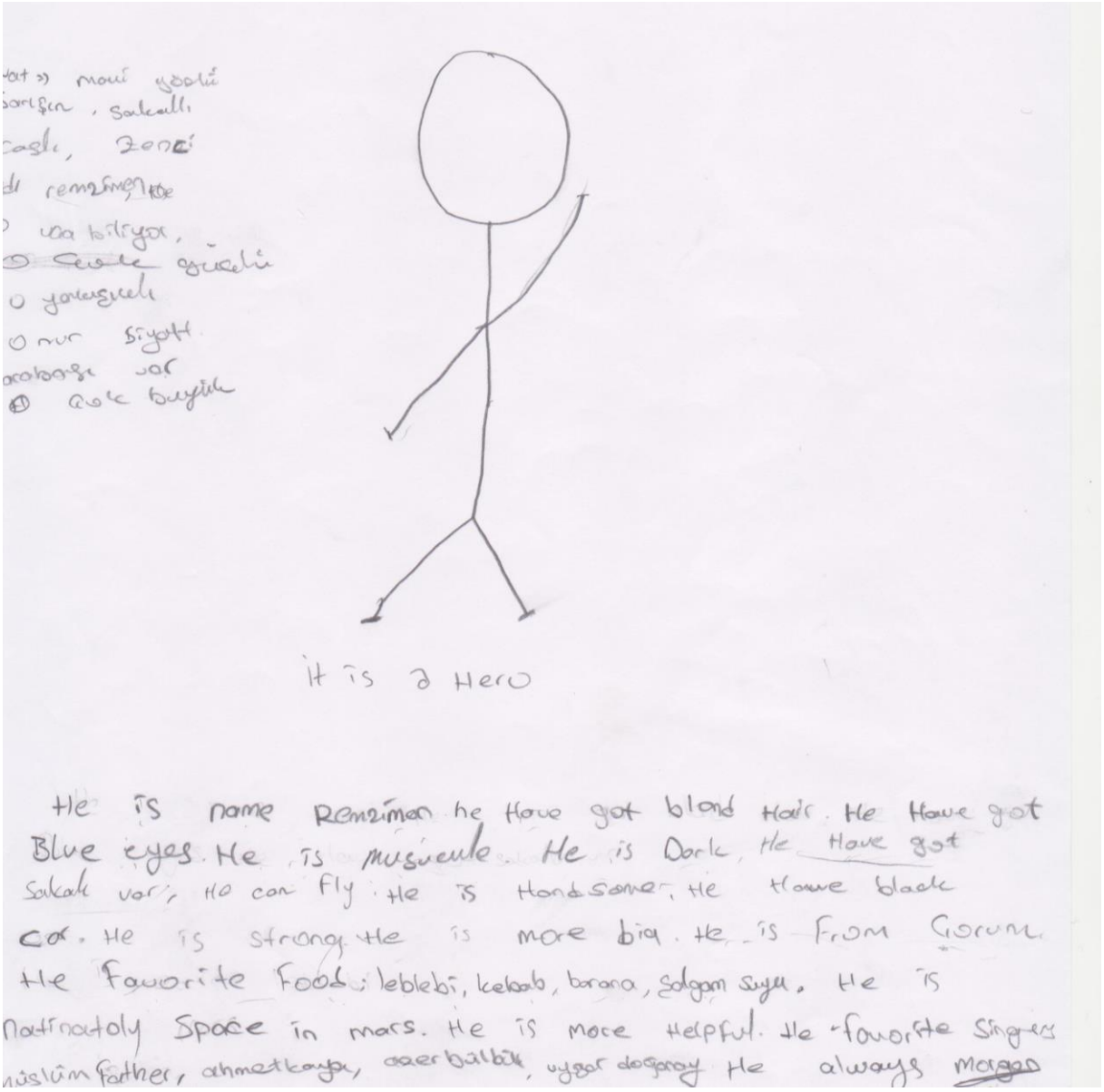
____ Yurt dışı seyahatlerimde işime yarayacak

____ Gelecekte İngilizce kullanmayı planlamıyorum.

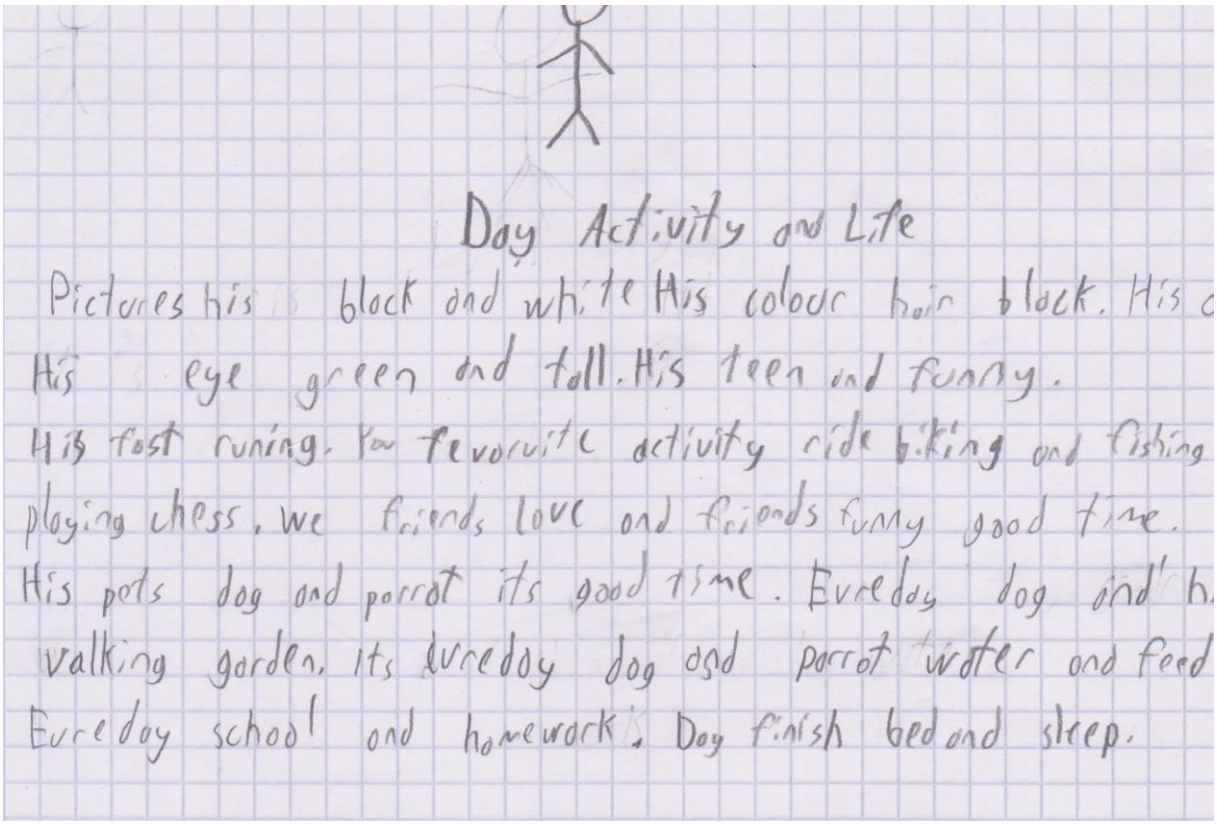
____ Diğer (lütfen belirtiniz)

_____

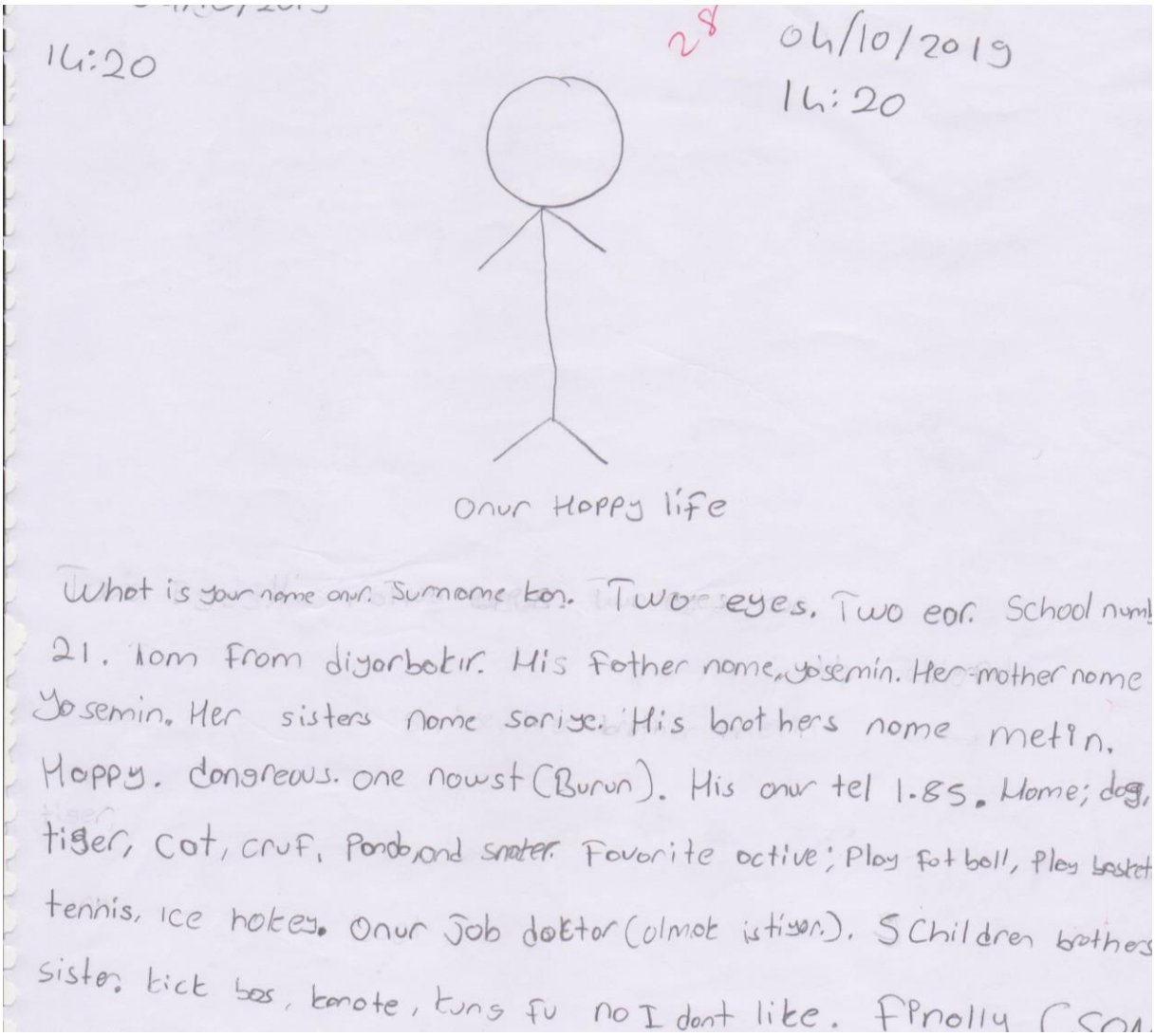
Appendix 2: Samples from the Products in the Pre-test



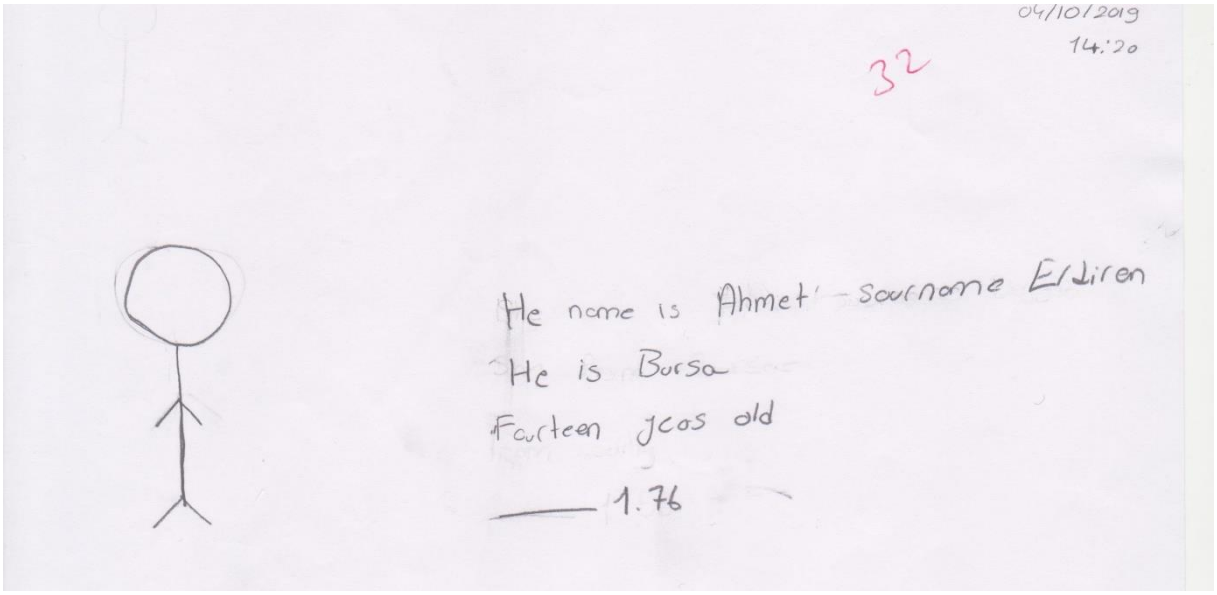
(The products belongs to S10.)



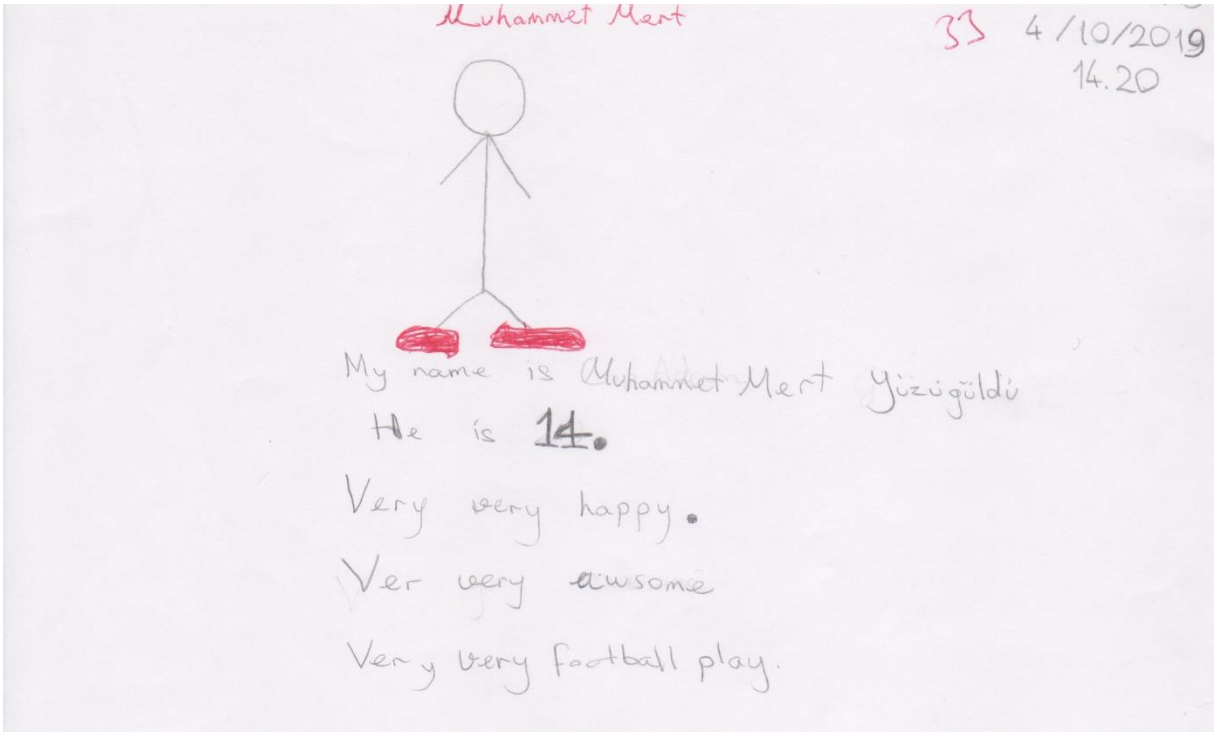
(The product belongs to S30.)



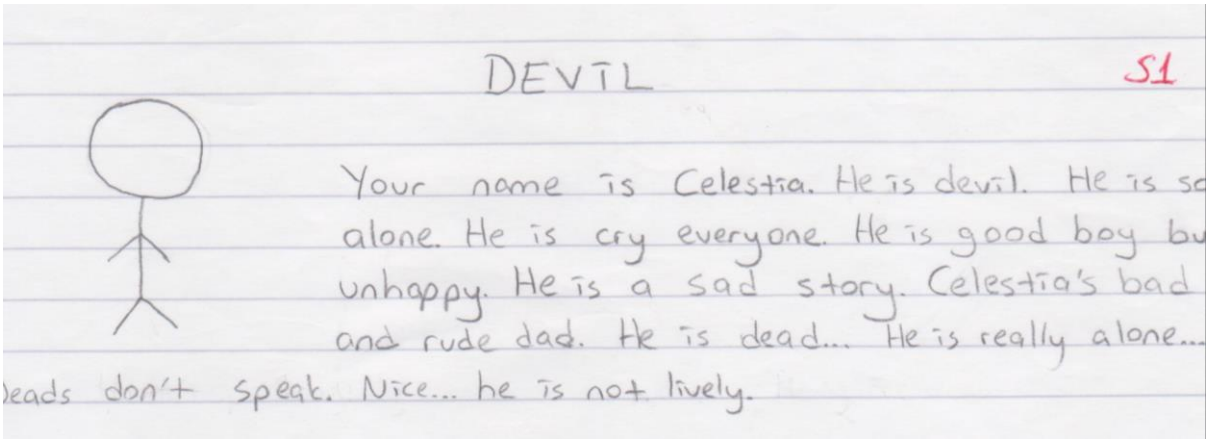
(The products belongs to S28.)



(The products belongs to S32.)



(The products belongs to S33.)



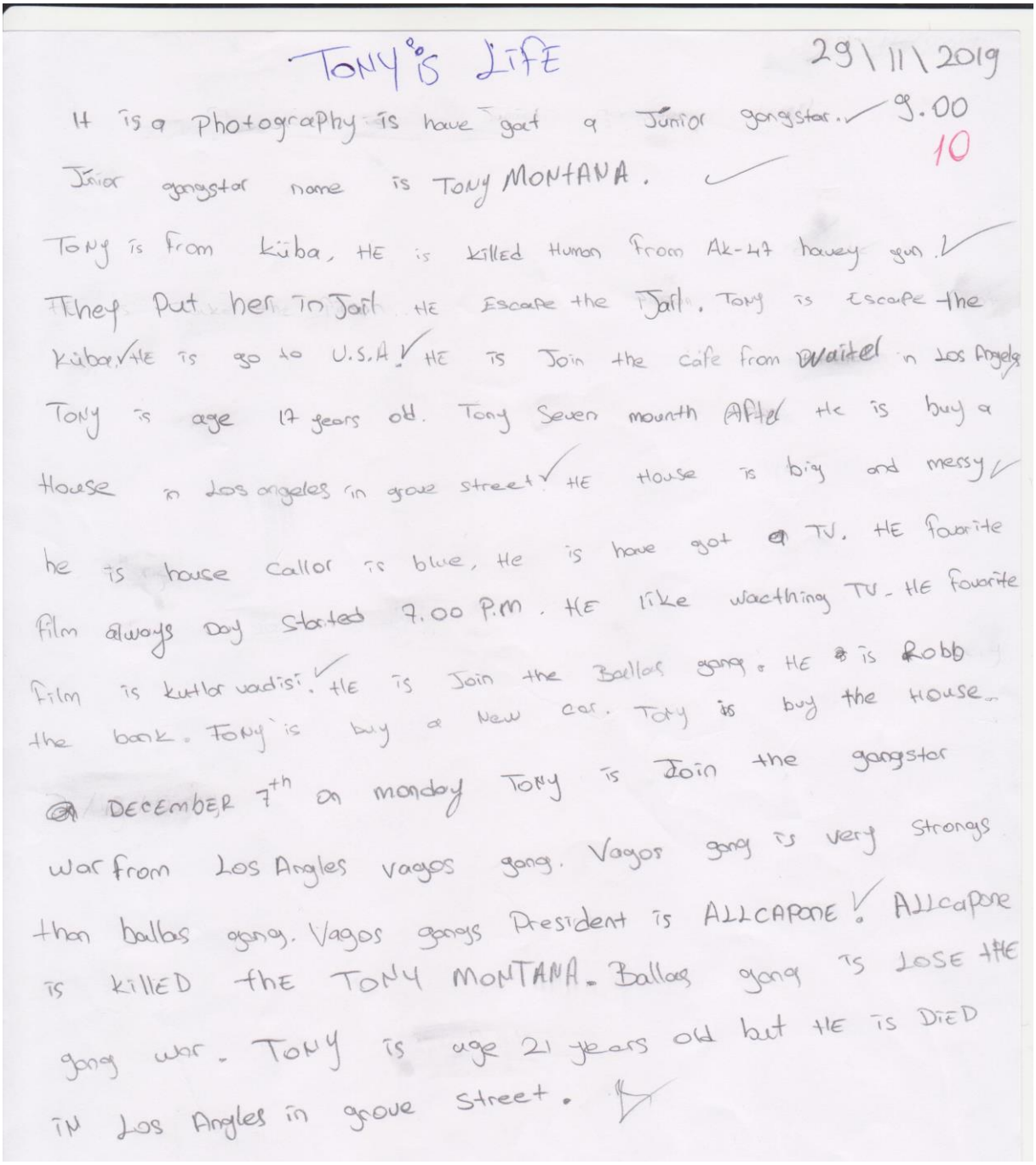
(The products belongs to S1.)

Appendix 3: The Prompt Image for the Post-test

The image is adopted from a doctorate thesis by Uysal (2009).

Appendix 4: Samples from the Products in the Post-test

The first part of their writing was written without getting help from GT.



TONY'S LIFE 29/11/2019

It is a Photography is have got a Junior gongstar. 9.00

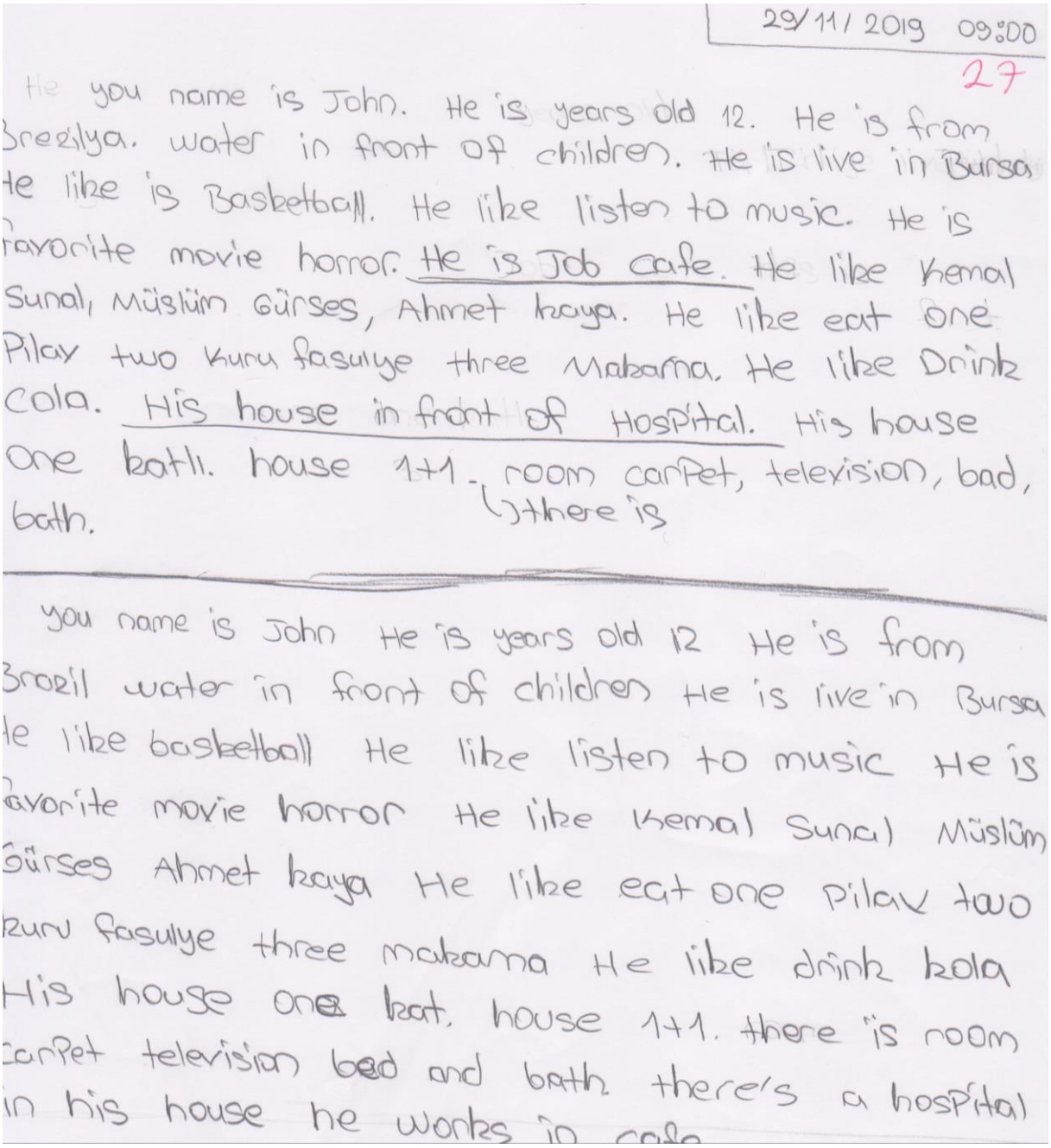
10

Jnior gongstar name is TONY MONTANA.

TONY is from kuba, HE is killed Human from Ak-47 havey gun. They put her in Jail HE Escape the Jail. Tony is escape the kuba. HE is go to U.S.A. HE is Join the cafe from waiter in Los Angels Tony is age 17 years old. Tony Seven mounth After he is buy a House in Los angeles in grove street. HE House is big and messy he is house callor is blue, He is have got a TV. HE fouarite film always Day Started 7.00 P.M. HE like wacthing TV. HE fouarite film is kutlar vadisi. HE is Join the Ballas gong. HE is Robb the bank. TONY is buy a New car. Tony is buy the House.

DECEMBER 7th on monday TONY is Join the gongstar war from Los Angles vagos gong. Vagos gong is very Strongs than ballas gong. Vagos gongs President is ALLCAPONE. Allcapone is killED the TONY MONTANA. Ballas gong is LOSE the gong war. TONY is uge 21 years old but HE is DIED IN Los Angles in groue street.

(The product belongs to S10.)



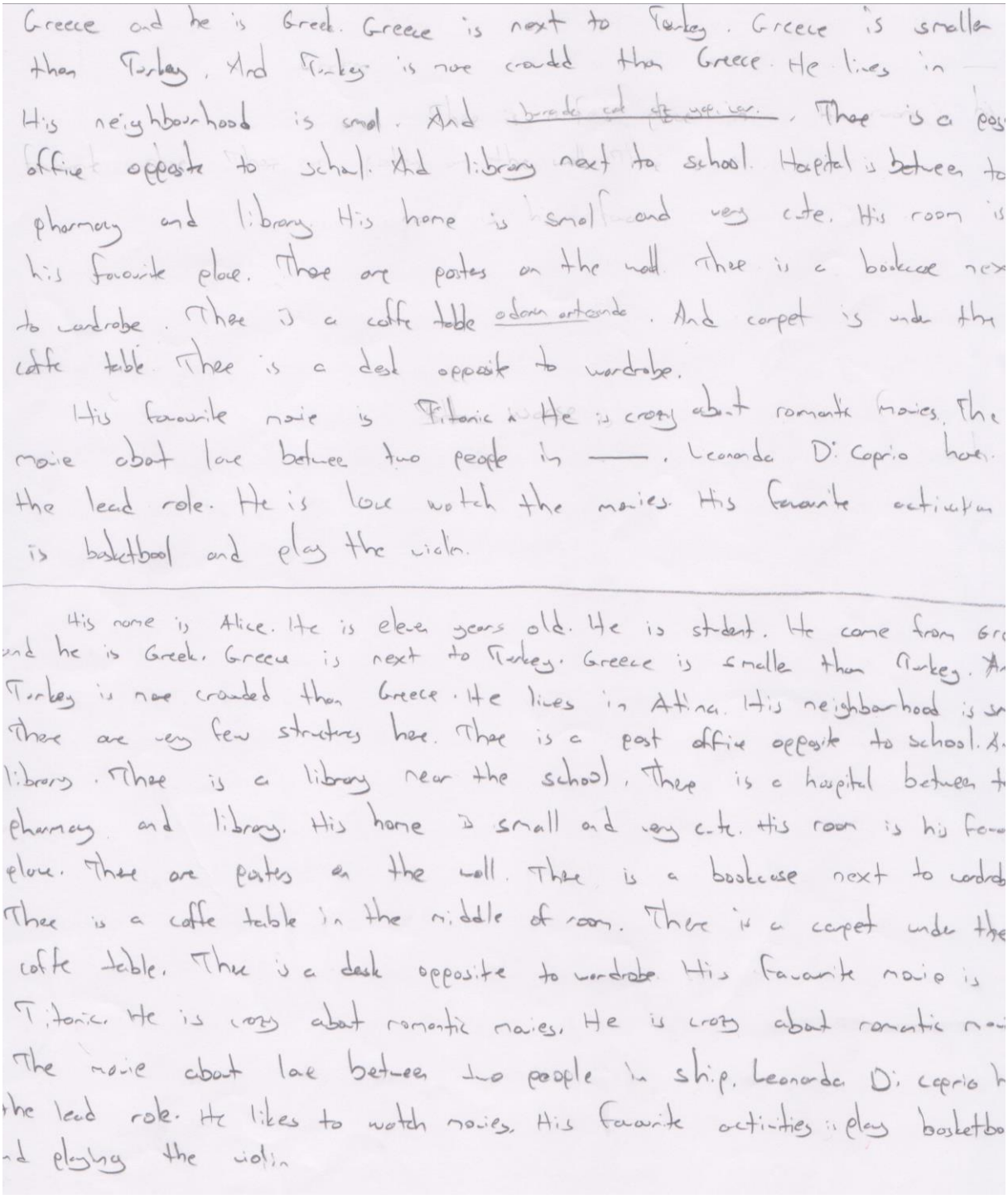
29/11/2019 09:00

27

He you name is John. He is years old 12. He is from Brezilya. water in front of children. He is live in Bursa He like is Basketball. He like listen to music. He is favorite movie horror. He is Job cafe. He like Kemal Sunal, Müslüm Gürses, Ahmet kaya. He like eat one Pilav two kuru fasulye three Makarna. He like Drink Cola. His house in front of Hospital. His house one bath. house 1+1. room carpet, television, bad, bath. there is

you name is John He is years old 12 He is from Brazil water in front of children He is live in Bursa He like basketball He like listen to music He is favorite movie horror He like Kemal Sunal Müslüm Gürses Ahmet kaya He like eat one Pilav two kuru fasulye three makarna He like drink kola His house one bat. house 1+1. there is room carpet television bed and bath there's a hospital in his house he works in cafe

(The product belongs to S27.)

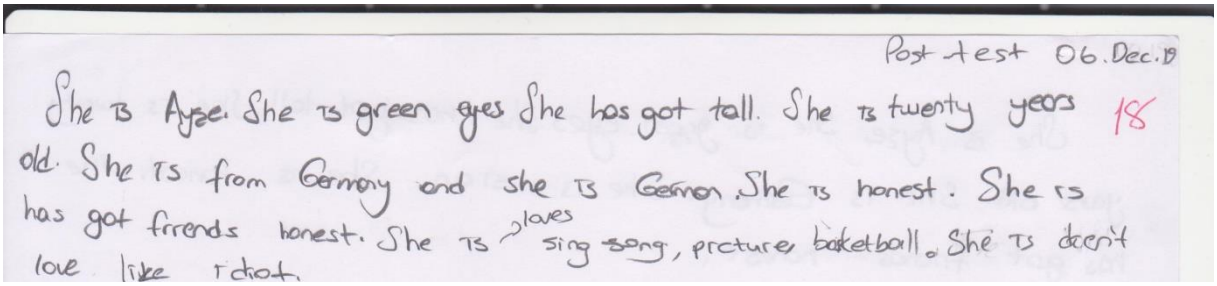


Greece and he is Greek. Greece is next to Turkey. Greece is smaller than Turkey. And Turkey is more crowded than Greece. He lives in His neighbourhood is small. And [illegible]. There is a post office opposite to school. And library next to school. Hospital is between to pharmacy and library. His home is small and very cute. His room is his favourite place. There are posters on the wall. There is a bookcase next to wardrobe. There is a coffe table odorn ortasında. And carpet is under the coffe table. There is a desk opposit to wardrobe.

His favourite movie is Titanic. He is crazy about romantic movies. The movie about love between two people in ——. Leonardo Di Caprio have the lead role. He is love watch the movies. His favourite activities is basketbool and play the violin.

His name is Alice. He is eleven years old. He is student. He come from Gr and he is Greek. Greece is next to Turkey. Greece is smaller than Turkey. A Turkey is more crowded than Greece. He lives in Atina. His neighborhood is s There are very few structures here. There is a post office opposite to school. A library. There is a library near the school. There is a hospital between t pharmacy and library. His home is small and very cute. His room is his fav place. There are posters on the wall. There is a bookcase next to wardrob There is a coffe table in the middle of room. There is a carpet under the coffe table. There is a desk opposite to wardrobe. His favourite movie is Titanic. He is crazy about romantic movies. He is crazy about romantic mo The movie about love between two people in ship. Leonardo Di Caprio h the lead role. He likes to watch movies. His favourite activities is play basketbo and playing the violin

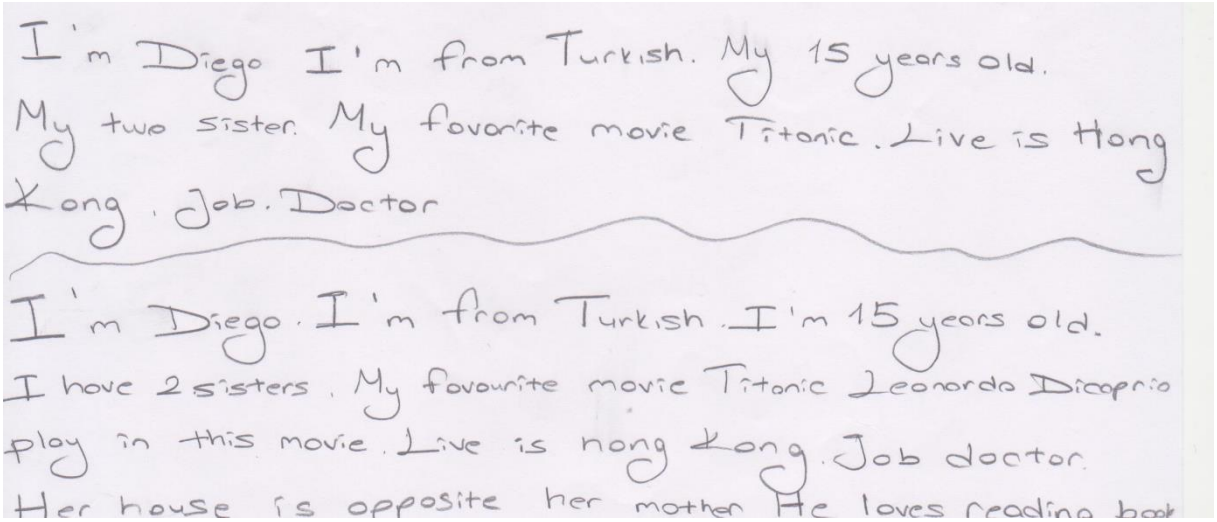
(The product belongs to S15.)



Post test 06.Dec.19

She is Ayse. She is green eyes She has got tall. She is twenty years old. She is from Germany and she is German. She is honest. She is has got friends honest. She is loves sing song, picture, baketball. She is doesn't love like riot. 18

(The product belongs to S18.)


I'm Diego I'm from Turkish. My 15 years old.
My two sister. My favorite movie Titonic. Live is Hong
Kong. Job. Doctor

I'm Diego. I'm from Turkish. I'm 15 years old.
I have 2 sisters. My favourite movie Titonic Leonardo Dicoprio
play in this movie. Live is nong Kong. Job doctor.
Her house is opposite her mother He loves reading book

(The product belongs to S31.)

Appendix 5: Traning Session

Training Session – Lesson Outline

I) Online look-up tools for foreign languages

A) Online dictionaries

1) Advantages

• Free of charge
• Easy to use: enter a word and voilà!
• Often several definitions, examples for each word or phrase to help fit the context

2) Disadvantages

• Have to pick the right word/phrase
• Not as complete as most full-sized dictionaries
• Doesn't do conjugation, agreement, etc. for you

B) Online translators

1) Advantages
• Free of charge
• Easy to use: enter word, phrase, sentence, paragraph...and voilà!
• Nothing to choose from, only one word/phrase comes out
• Attempts to do agreement, conjugation, etc.

2) Disadvantages

• No options you can choose from: it decides for you
• Errors in word choice, agreement
• Unable to detect context, subject matter
• Against the rules for many other departments/instructors

II) Native-language judgments

1) Easier to understand and spot errors in your native language than in a foreign language.
2) Click on Google Translate application
3) Set translator to "English to Turkish"
4) Enter this poem:
I must go seek some dewdrops here,
And hang a pearl in every cowslip's ear. (Shakespeare, retrieved from https://www.williamshakespeare.net/a-fairy-song.jsp on 12 Sep 2019)
5) What is the result you get in English? Do you see any mistakes?

_____

(Buraya biraz çiy damlası aramalıyım ve her inek için bir inci asmalıyım.)
6) here: very common word in English, but when translated into Turkish its conjugation depends on the context. Apparently Google Translate cannot get the context.

Cowslip: a small European wild plant with sweet smelling yellow flowers (Longman Dictionary of Contemporary English, 1995)
The Google Translate makes a mistake in word choice.
7) Enter "Give me a hand with that."
8) What is the result you get in English? Do you see any mistakes?

_____

(Bana yardım et.)
How would you want help from somebody in Turkish?

_____

Is this translations satisfactory to you? Why or why not?

III) Foreign-language judgments

E.g 1:

1) As a non-native speaker of English, you may be able to notice some mistakes or be able to understand even if there is a mistake, but harder than in your native language.
2) Click on the Google Translate application
3) Set translator to "Turkish to English"
4) Enter "Şu işe bir el atsana."
5) What is the result you get in English? Do you see any mistakes?

_____

(Give me a hand with that.)
6) You can see the application cannot get the daily language and translates the same expressions in different ways.

E.g 2:

1) Set translator to "Turkish to English"
2) Enter "Babamın arabası sarı."
3) What is the result you get in English? Do you see any mistakes?

_____

(My father's car is yellow.)
4) Google Translate gave a correct translation.
5) Enter "Babamın saçları sarı."
6) What is the result you get in English? Do you see any mistakes?

_____

(My father's hair is yellow.)
7) Enter "Babam sarı saçlara sahip."
8) What is the result you get in English? Do you see any mistakes?

_____

(Dad has got blond hair.)
17) Google Translate offers different translations for the same word "sarı" in these sentences. Which one is true? Why do you think there is a difference?
18) Although we use blond to express the colour of hair in English, it has the same equivalence in Turkish. Be careful about such words.

IV) Example translations from Turkish to English

Now we're going to look at some translations from Turkish to English to see which ones are correct, could be correct depending on context, or are incorrect (at least partially).

A) Enter the following sentences into Google Translate, selecting "Turkish to English".

1) Evini hiç beğenmedim gerçekten. (I didn't really like your house)
• Correct / Could be correct / Incorrect?
• Corrected sentence/Other options?

2) Onun evini gerçekten hiç beğenmedim. (I really didn't like his house.)
• Correct / Could be correct / Incorrect?
• Corrected sentence/Other options?

3) Sözlükten bakar mısın? (Can you look at the dictionary.)

• Correct / Could be correct / Incorrect?
• Phrasal verbs in English (e.g. to look up, to sit down, etc.) don't exist in Turkish
• Corrected form?

4) Uyandı. (She woke up.)
• Correct / Could be correct / Incorrect?
• Phrasal verbs in English (e.g. to wake up, to sit down, etc.) don't exist in Turkish.
• Uyanmak has some connotations in Turkish, which is not shared in English, such as getting aware of something after a long time. Here GT cannot such a colloquial expression.
• Corrected sentence/Other options?

5) Naber? (What's up?)
• Correct / Could be correct / Incorrect?
• Familiar/colloquial expressions in English and Turkish
• Corrected sentence/Other options?

6) Eline sağlık. (Health in hand)
• Correct / Could be correct / Incorrect?
• Different cultural expressions in English and Turkish
• Corrected sentence/Other options?

7) Arkadaşım Amerikalı. (My friend American)
• Correct / Could be correct / Incorrect?
• Corrected sentence/Other options?

8) Arkadaşım Amerikalıdır. (My is American.)
• Correct / Could be correct / Incorrect?
• Different tenses in French and English
• Corrected sentence/Other options?

9) Dedem 70'lerin sonlarında. (My grandfather in the late seventies.)
• Correct / Could be correct / Incorrect?
• Corrected sentence/Other options?

10) Dedem 70'lerin sonlarındadır. (My grandfather is in the late seventies.)
• Correct / Could be correct / Incorrect?
• Corrected sentence/Other options?

B) Enter the following sentences into Google Translate, selecting "English to Turkish".

1) I'm never wrong. (I am never wrong.)
• Correct / Could be correct / Incorrect?
• Corrected sentence/Other options?

2) Got to go. (Gitmeliyim.)
• Correct / Could be correct / Incorrect?
• "Incorrect" grammar in English

• Corrected sentence/Other options?

3) Ali rode a horse. (Ali ata bindi.)
• Correct / Could be correct / Incorrect?

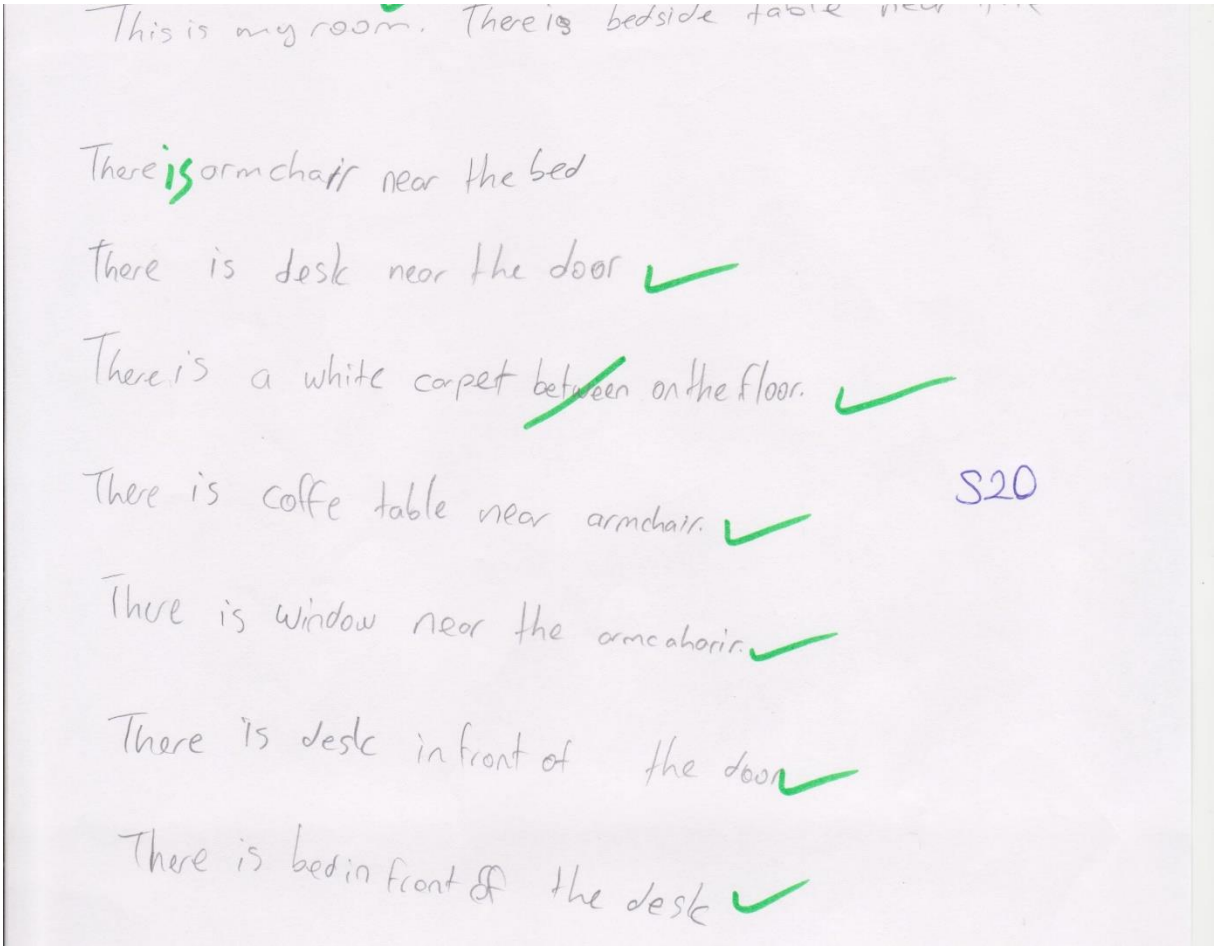
4) Ancestor (ata)
• Correct / Could be correct / Incorrect?
• Conjugations and homonyms in Turkish
• Corrected sentence/Other options?

5) In the garden there was a child holding a bag in his hand. (Bahçede elinde çanta tutan bir çocuk vardı.)
• Correct / Could be correct / Incorrect?
• Feminine and masculine words in English
• Corrected sentence/Other options?

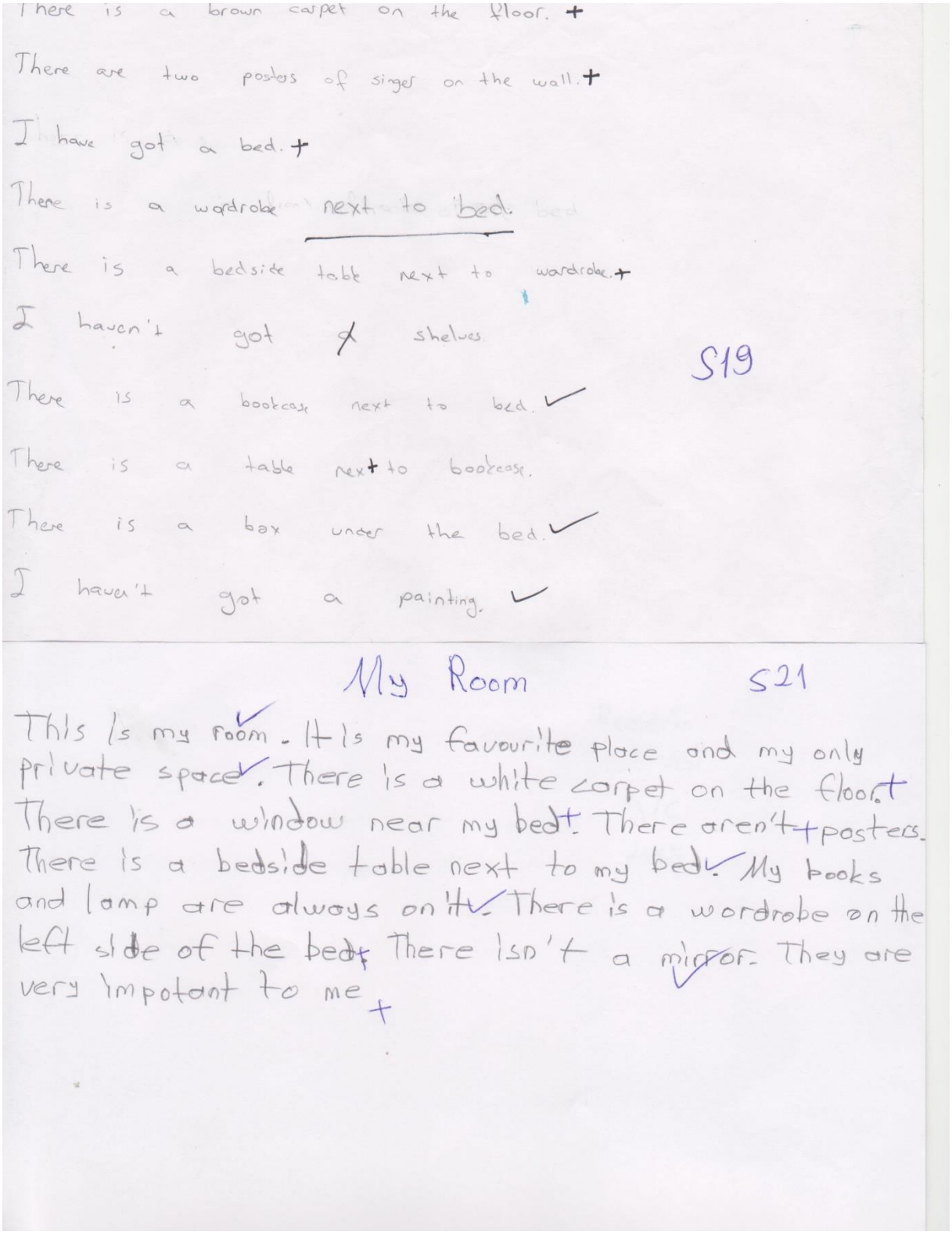
6) a television (bir televizyon)
• Correct / Could be correct / Incorrect?
• Isolated words in Turkish and English
• Corrected sentence/Other options?

7) set (gives a long list)
• Correct / Could be correct / Incorrect?
• Isolated words in Turkish and English

Appendix 6: Samples from the First Week Activity



(The product belongs to S20)



There is a brown carpet on the floor. +

There are two posters of singer on the wall. +

I have got a bed. +

There is a wardrobe next to bed.

There is a bedside table next to wardrobe. +

I haven't got ~~a~~ shelves.

S19

There is a bookcase next to bed. ✓

There is a table next to bookcase.

There is a box under the bed. ✓

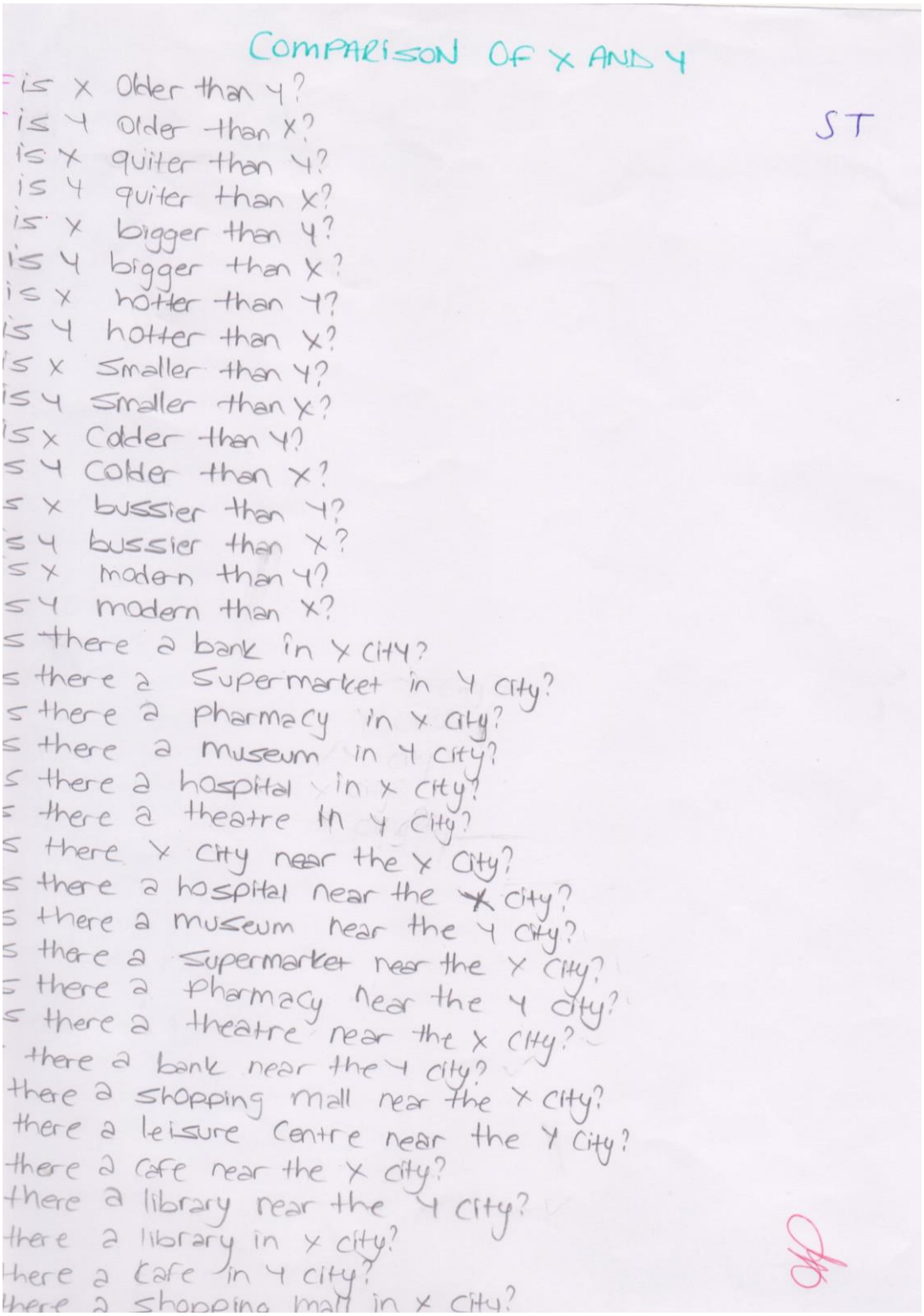
I haven't got a painting. ✓

My Room S21

This is my room. It is my favourite place and my only private space. There is a white carpet on the floor. + There is a window near my bed. There aren't + posters. There is a bedside table next to my bed. ✓ My books and lamp are always on it. ✓ There is a wordrobe on the left side of the bed. There isn't a mirror. ✓ They are very impotant to me. +

(The products belong to S19 and S21)

Appendix 7: Samples from the Second Week Activity

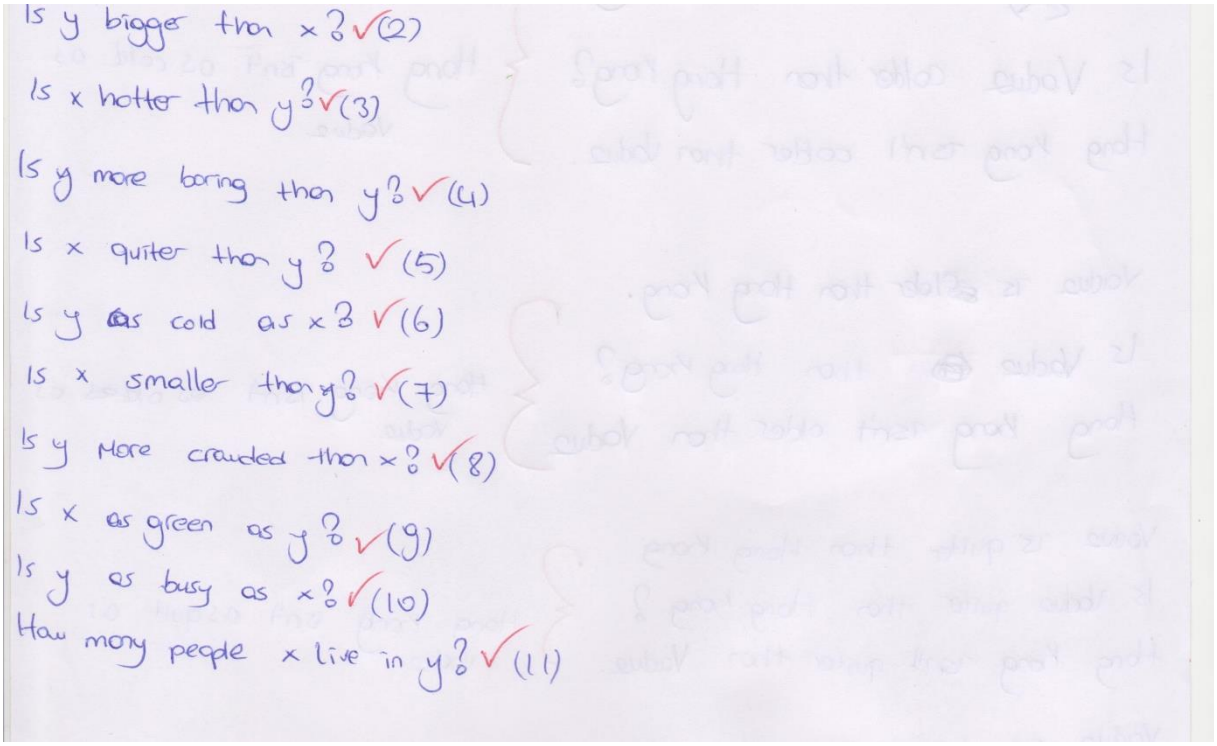


COMPARISON OF X AND Y

ST

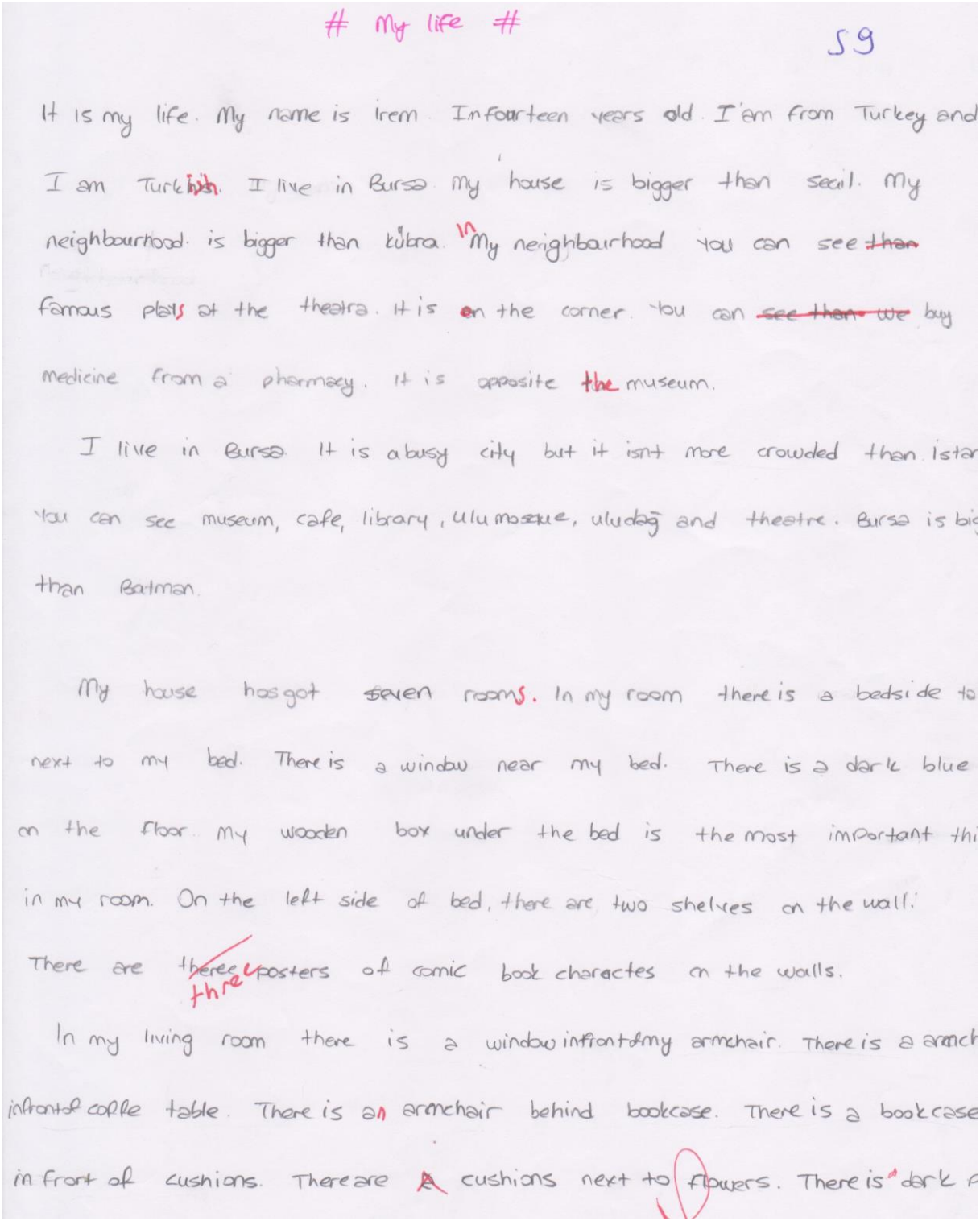
is x older than y?
is y older than x?
is x quiter than y?
is y quiter than x?
is x bigger than y?
is y bigger than x?
is x hotter than y?
is y hotter than x?
is x smaller than y?
is y smaller than x?
is x colder than y?
s y colder than x?
s x bussier than y?
s y bussier than x?
s x modern than y?
s y modern than x?
s there a bank in x city?
s there a supermarket in y city?
s there a pharmacy in x city?
s there a museum in y city?
s there a hospital in x city?
s there a theatre in y city?
s there x city near the x city?
s there a hospital near the ~~x~~ city?
s there a museum near the y city?
s there a supermarket near the x city?
s there a pharmacy near the y city?
s there a theatre near the x city?
there a bank near the y city?
there a shopping mall near the x city?
there a leisure centre near the y city?
there a cafe near the x city?
there a library near the y city?
there a library in x city?
there a cafe in y city?
there a shopping mall in x city?

(The product belongs to S7.)



(The product belongs to S18)

Appendix 8: Samples from the Third Week Activity



My life

S9

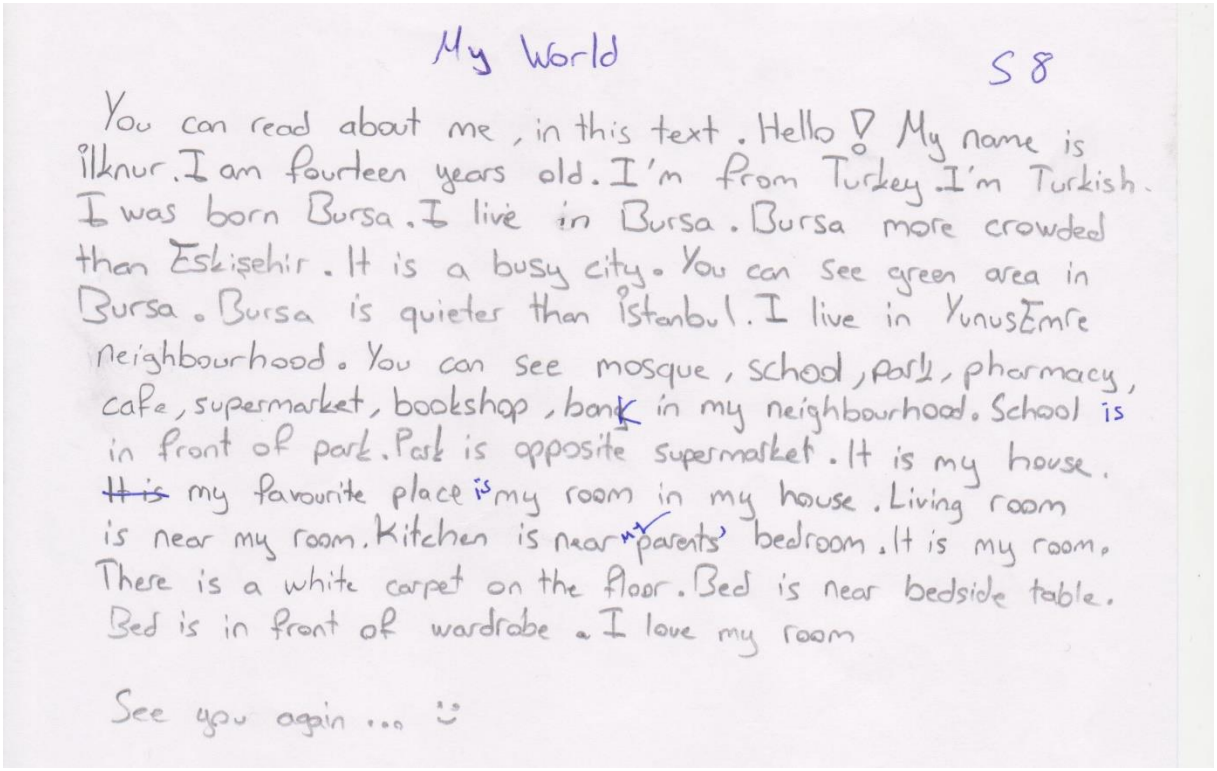
It is my life. My name is Irem. I'm fourteen years old. I am from Turkey and I am Turkish. I live in Bursa. My house is bigger than Seul. My neighbourhood is bigger than Kübra. In My neighbourhood you can see ~~than~~ famous plays at the theatre. It is on the corner. You can ~~see then we~~ buy medicine from a pharmacy. It is opposite the museum.

I live in Bursa. It is a busy city but it isn't more crowded than Istar You can see museum, cafe, library, Ulu mosque, uludağ and theatre. Bursa is bi than Batman.

My house has got ~~seven~~ rooms. In my room there is a bedside to next to my bed. There is a window near my bed. There is a dark blue on the floor. My wooden box under the bed is the most important thi in my room. On the left side of bed, there are two shelves on the wall. There are ~~there~~ three posters of comic book characters on the walls.

In my living room there is a window infront of my armchair. There is a armch infront of coffe table. There is an armchair behind bookcase. There is a bookcase in front of cushions. There are ~~A~~ cushions next to flowers. There is dark

(The product belongs to S9.)



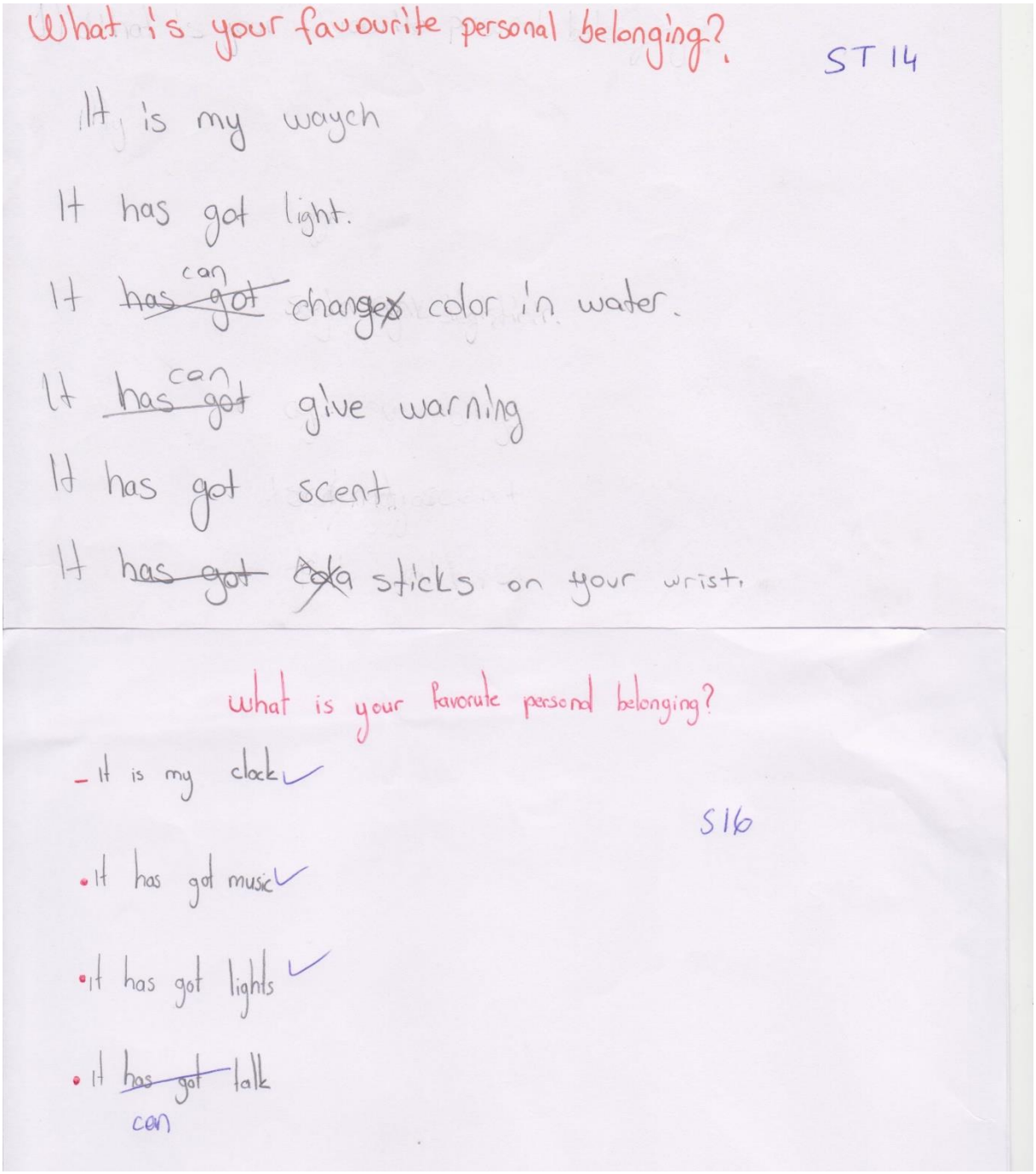
My World S8

You can read about me, in this text. Hello ☺ My name is İlknur. I am fourteen years old. I'm from Turkey. I'm Turkish. I was born Bursa. I live in Bursa. Bursa more crowded than Eskişehir. It is a busy city. You can see green area in Bursa. Bursa is quieter than İstanbul. I live in YunusEmre neighbourhood. You can see mosque, school, park, pharmacy, cafe, supermarket, bookshop, bank in my neighbourhood. School is in front of park. Park is opposite supermarket. It is my house. ~~It is~~ my favourite place is my room in my house. Living room is near my room. Kitchen is near my parents' bedroom. It is my room. There is a white carpet on the floor. Bed is near bedside table. Bed is in front of wardrobe. I love my room

See you again ... ☺

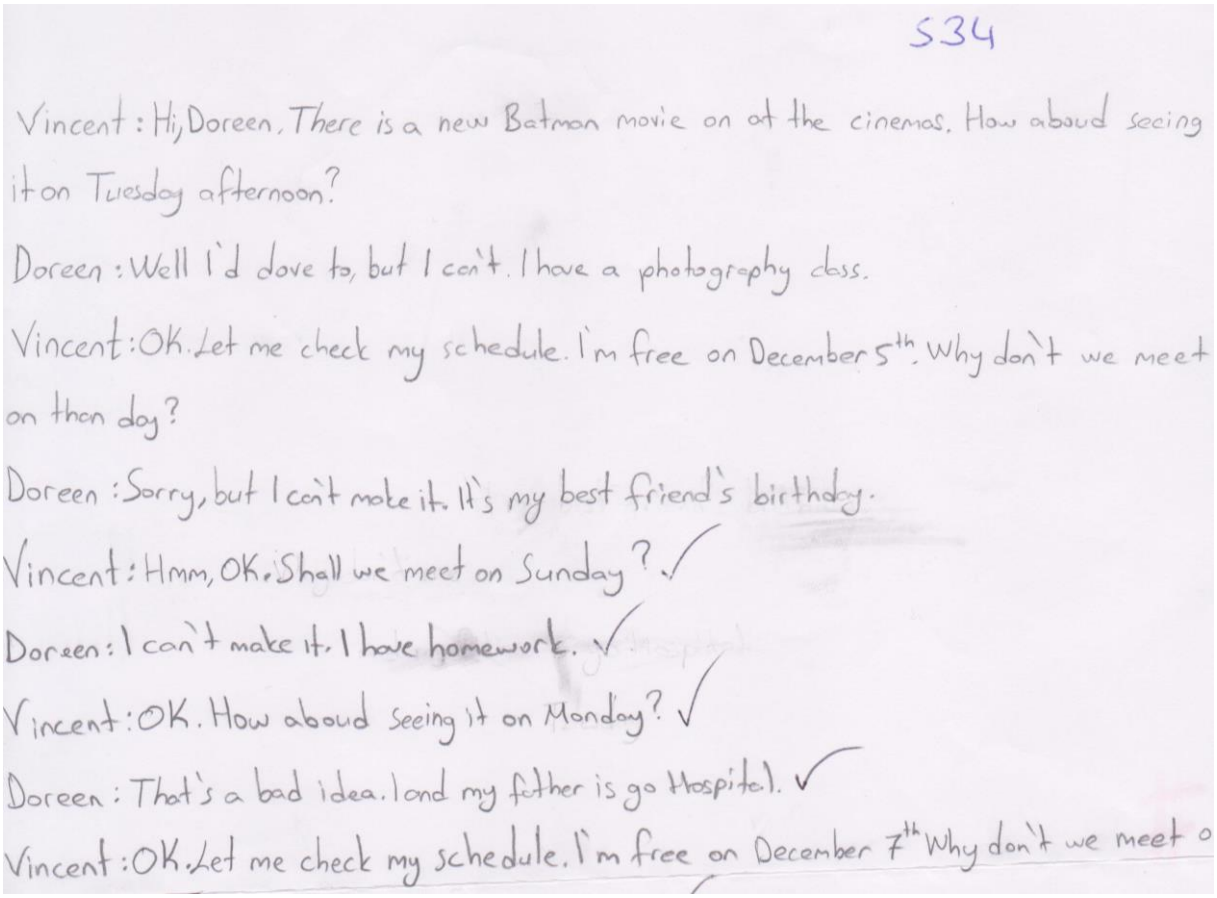
(The product belongs to S8.)

Appendix 9: Samples from the Forth Week Activity



(The products belong to S14 and S16.)

Appendix 10: Samples from the Fifth Week Activity


S34

Vincent: Hi, Doreen. There is a new Batman movie on at the cinemas. How about seeing it on Tuesday afternoon?

Doreen: Well I'd love to, but I can't. I have a photography class.

Vincent: OK. Let me check my schedule. I'm free on December 5th. Why don't we meet on than day?

Doreen: Sorry, but I can't make it. It's my best friend's birthday.

Vincent: Hmm, OK. Shall we meet on Sunday? ✓

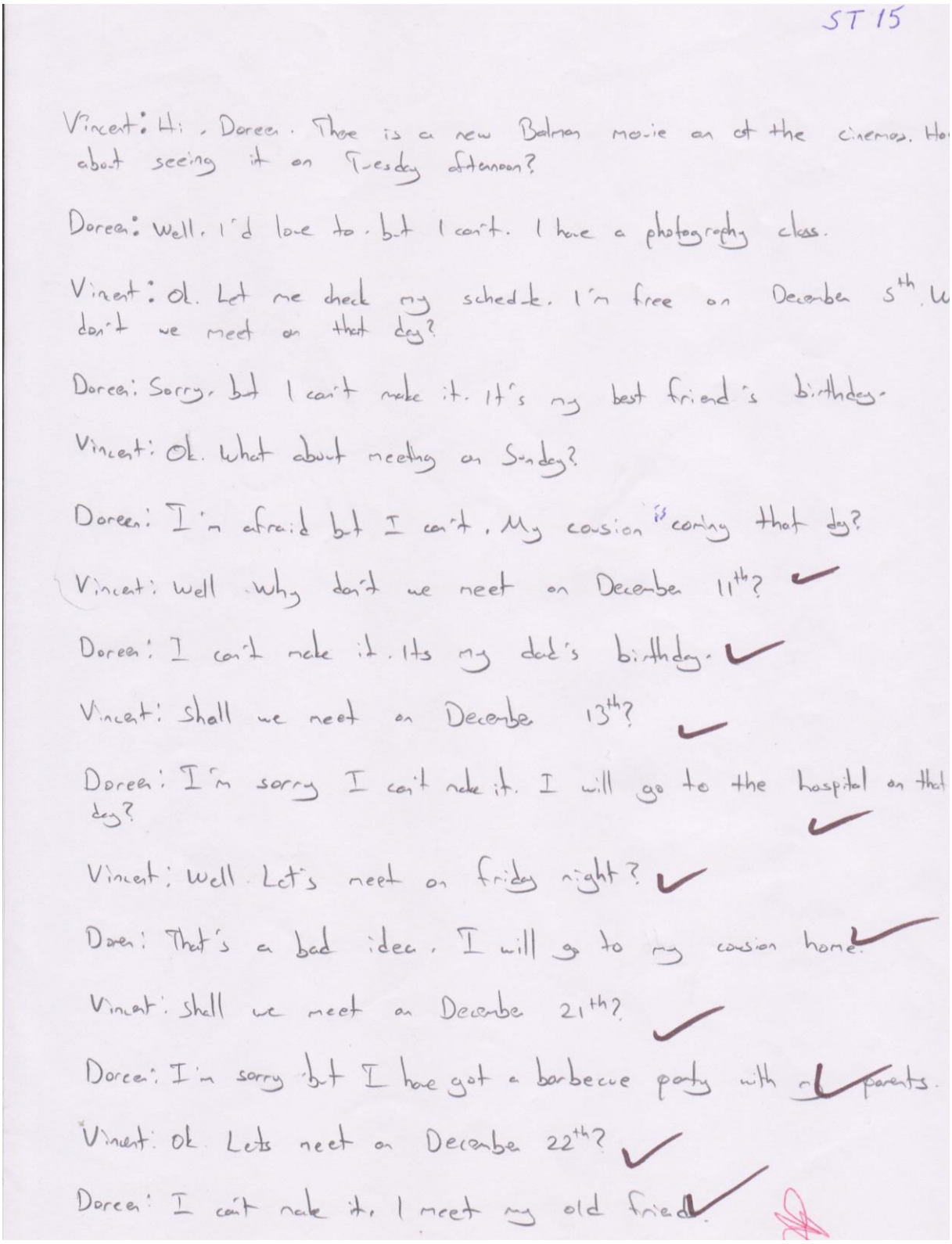
Doreen: I can't make it. I have homework. ✓

Vincent: OK. How about seeing it on Monday? ✓

Doreen: That's a bad idea. I and my father is go Hospital. ✓

Vincent: OK. Let me check my schedule. I'm free on December 7th Why don't we meet o

(The product belongs to S34.)



ST 15

Vincent: Hi, Doreen. There is a new Batman movie on at the cinemas. How about seeing it on Tuesday afternoon?

Doreen: Well, I'd love to, but I can't. I have a photography class.

Vincent: Ok. Let me check my schedule. I'm free on December 5th. When don't we meet on that day?

Doreen: Sorry, but I can't make it. It's my best friend's birthday.

Vincent: Ok. What about meeting on Sunday?

Doreen: I'm afraid but I can't. My cousin is coming that day?

Vincent: Well why don't we meet on December 11th?

Doreen: I can't make it. Its my dad's birthday.

Vincent: Shall we meet on December 13th?

Doreen: I'm sorry I can't make it. I will go to the hospital on that day?

Vincent: Well Let's meet on friday night?

Doreen: That's a bad idea. I will go to my cousin home.

Vincent: Shall we meet on December 21th?

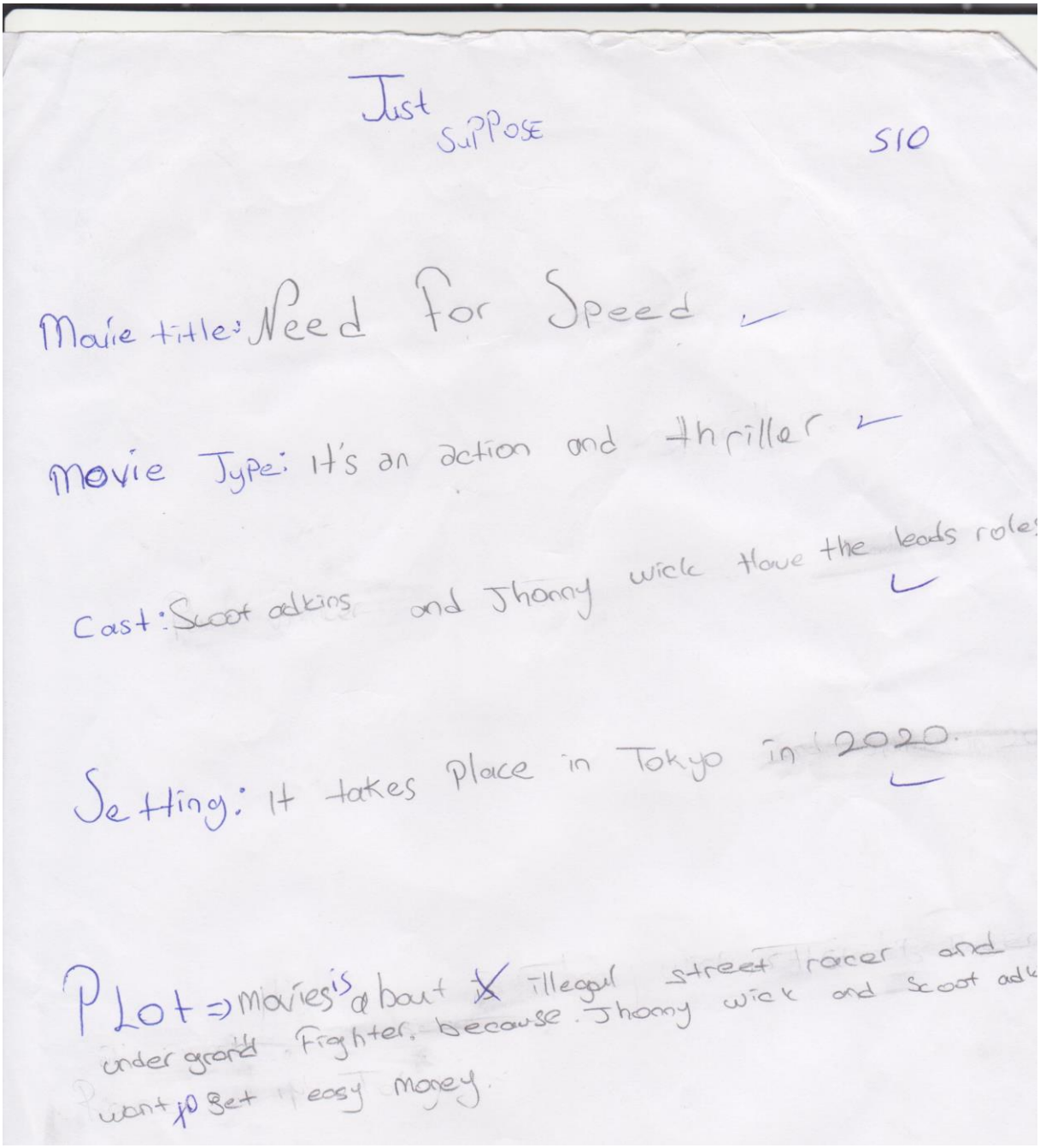
Doreen: I'm sorry but I have got a barbecue party with my parents.

Vincent: Ok Lets meet on December 22th?

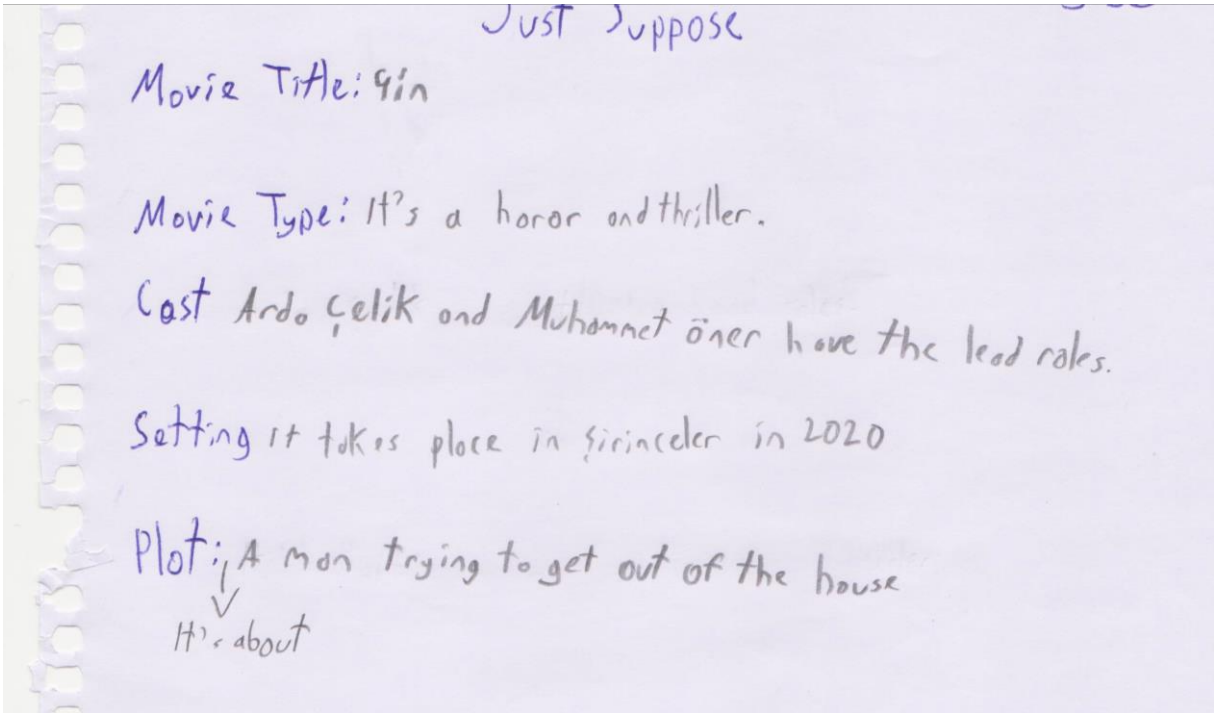
Doreen: I can't make it, I meet my old friend.

(The product belongs to S15.)

Appendix 11: Samples from the Sixth Week



(The product belong to S10.)


Just Suppose

Movie Title: Yin

Movie Type: It's a horor and thriller.

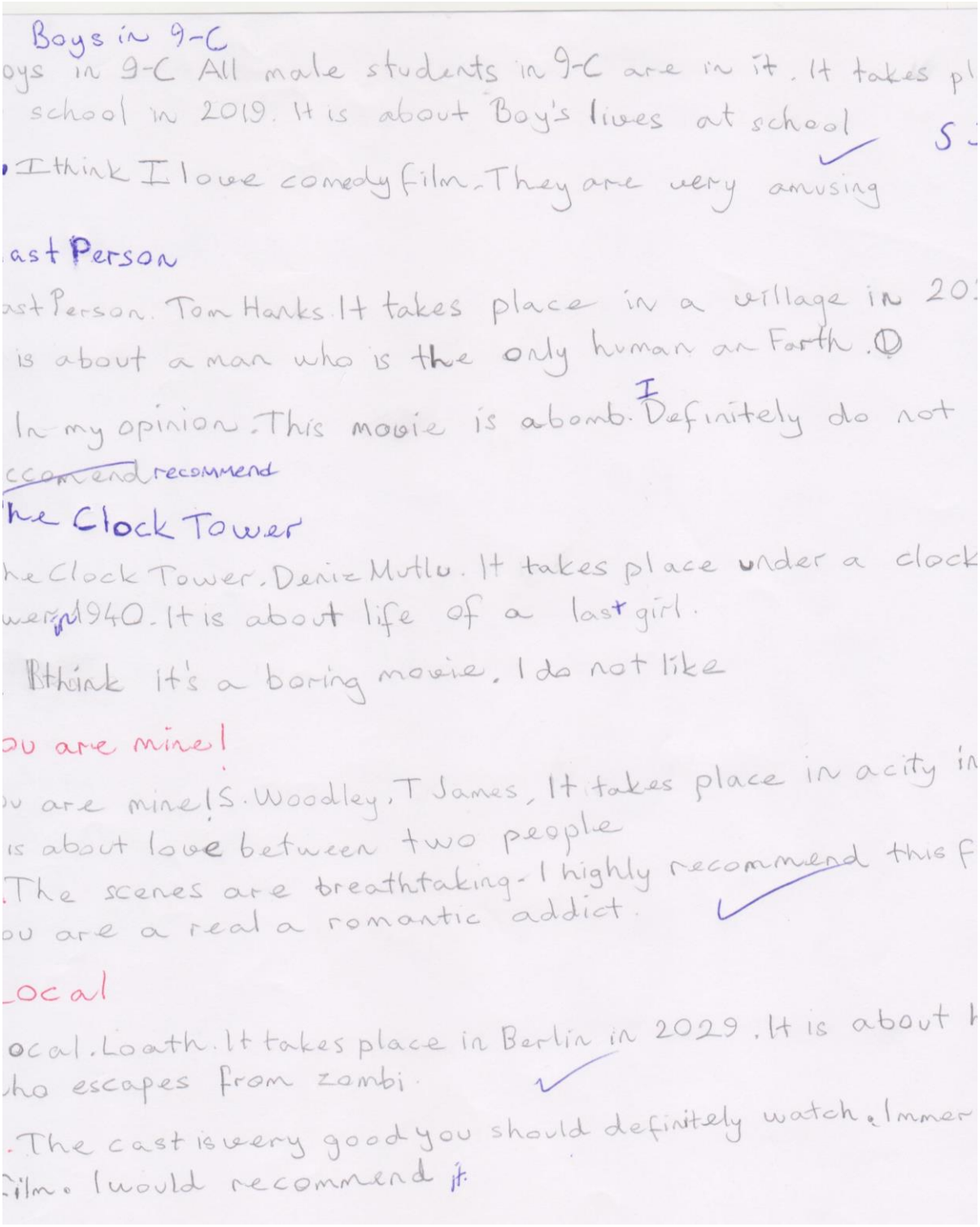
Cast Ardo Çelik and Muhammet öner have the lead roles.

Setting it takes place in Şirinceler in 2020

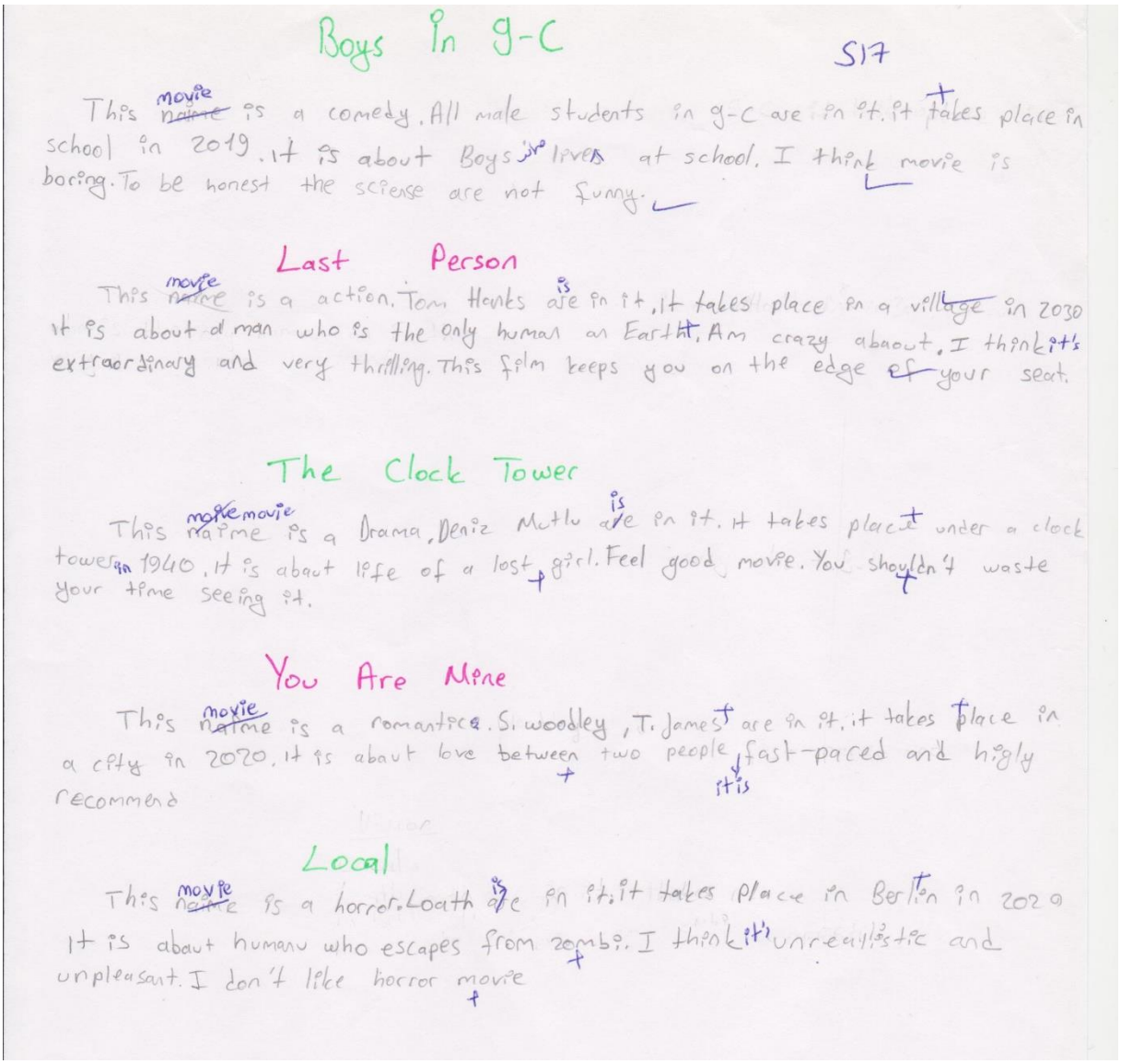
Plot: A man trying to get out of the house
↓
It's about

(The product belongs to S30.)

Appendix 12: Samples from the Seventh Week Activities



(The product belongs to S33.)



Boys In g-C

S17

This movie ~~name~~ is a comedy. All male students in g-C are in it. it takes place in school in 2019. it is about Boys ~~in~~ lives at school. I think movie is boring. To be honest the scense are not funny.

Last Person

This movie ~~name~~ is a action. Tom Hanks are in it, it takes place in a ~~village~~ in 2030 it is about a man who is the only human on Earth. Am crazy about. I think it's extraordinary and very thrilling. This film keeps you on the edge ~~of~~ your seat.

The Clock Tower

This ~~movie~~ movie ~~name~~ is a Drama, Deniz Mutlu are in it. It takes place under a clock tower in 1940, it is about life of a lost girl. Feel good movie. You shouldn't waste your time seeing it.

You Are Mine

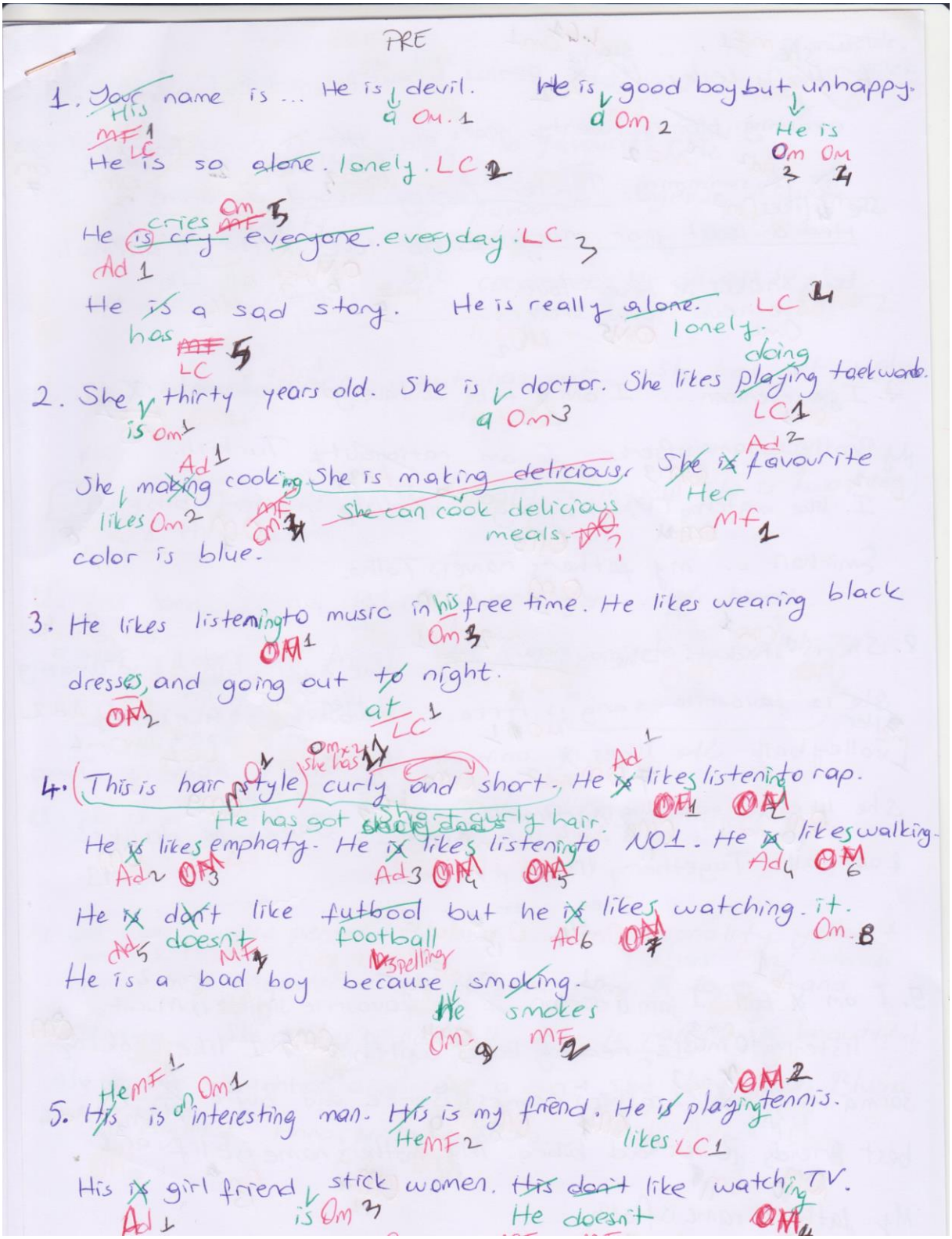
This movie ~~name~~ is a romantica. S. woodley, T. James are in it, it takes place in a city in 2020, it is about love between two people fast-paced and higly recommend

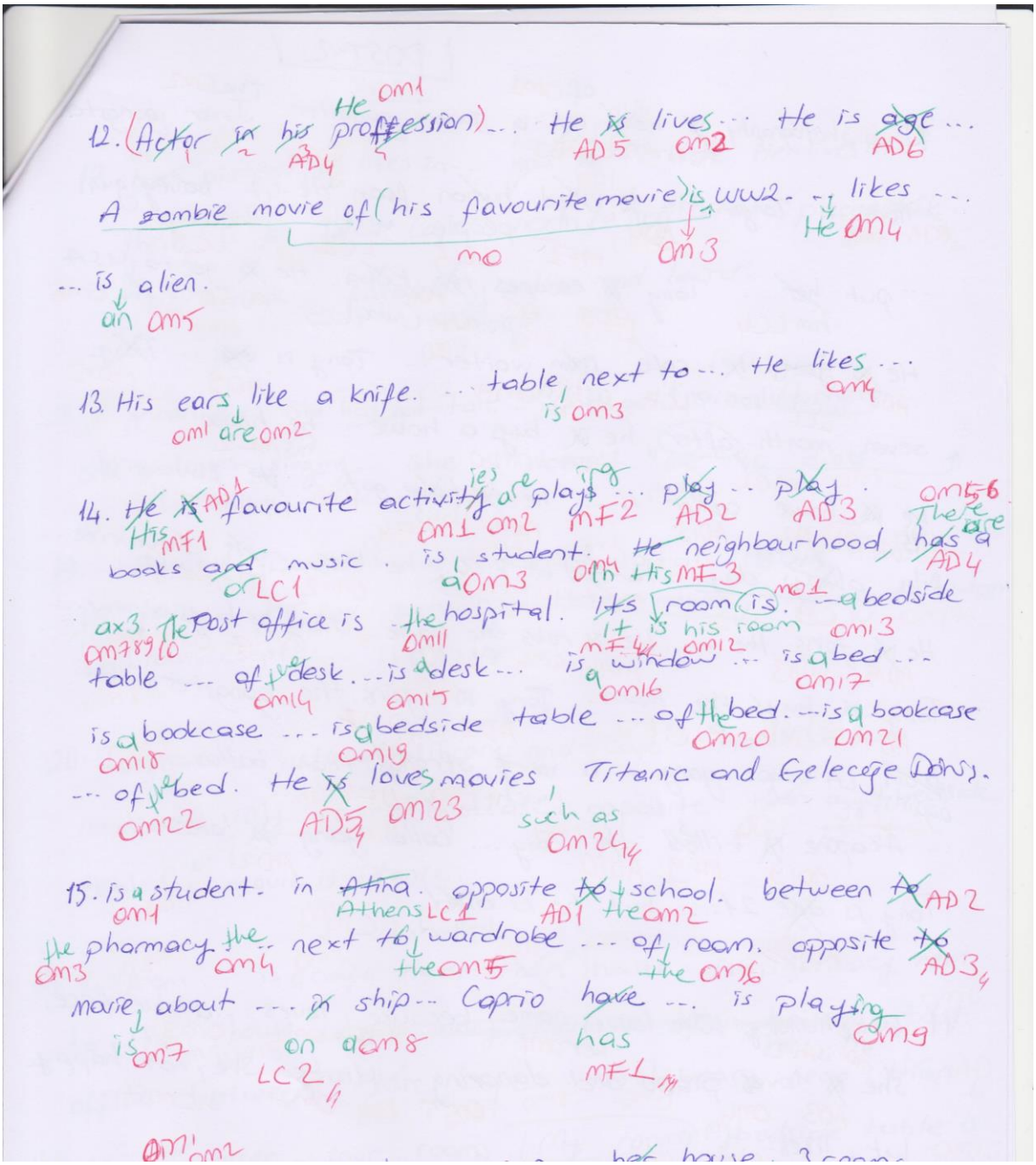
Local

This movie ~~name~~ is a horror. Loath are in it, it takes place in Berlin in 2020 it is about humans who escapes from zombi. I think it's unrealistic and unpleasant. I don't like horror movie

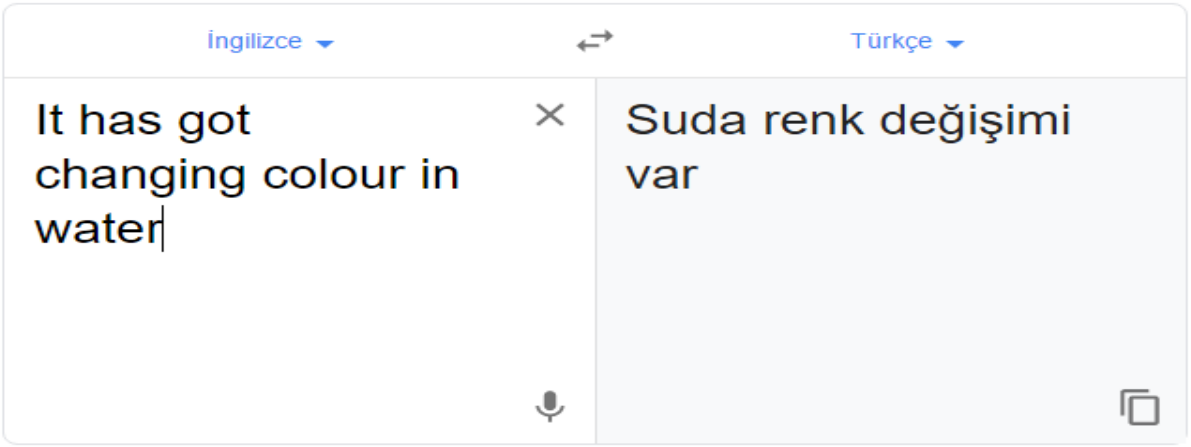
(The product belongs to S17.)

Appendix 13: Analysis of Grammatical and Lexical Errors in Pre and Post-tests

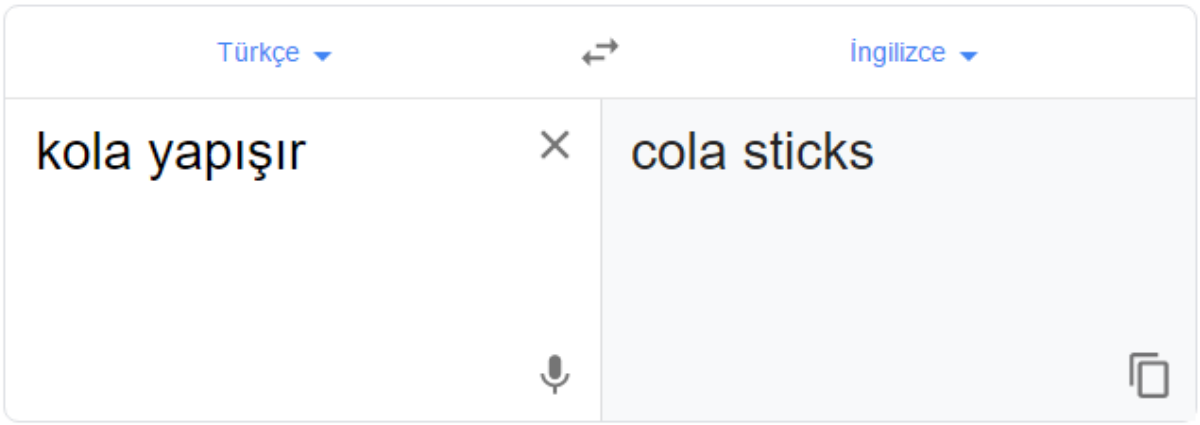




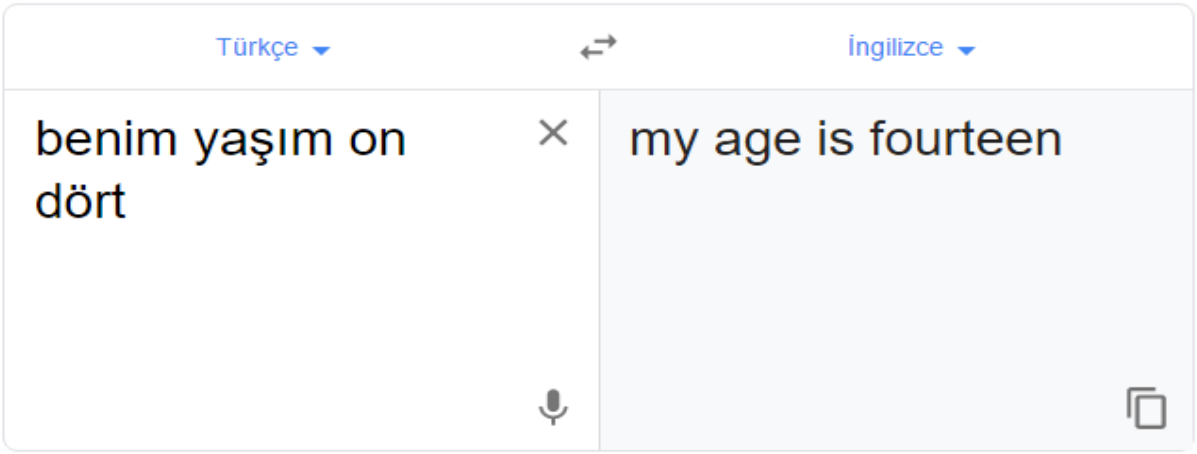
Appendix 14: Samples from Mistranslation on Google Translate



(The image was taken on October 31, 2019.)



(The image was taken on October 31, 2019.)



(The image was taken on October 25, 2019.)

Appendix 15: Bursa Provincial Directorate of National Education Approval Letter

T.C.
BURSA VALİLİĞİ
İl Millî Eğitim Müdürlüğü

Sayı : 86896125-605.01-E.14640661 07.08.2019
Konu : Ayşe TUZCU'nun Araştırma İzni

MÜDÜRLÜK MAKAMINA

İlgi : Milli Eğitim Bakanlığı'nın Araştırma, Yarışma ve Sosyal Etkinlik İzinleri konulu 22/08/2017 tarihli ve 2017/25 sayılı Genelgesi.

Uludağ Üniversitesi İngiliz Dili Eğitimi bölümü yüksek lisans öğrencisi Ayşe TUZCU'nun "Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi" konulu araştırma isteği Ayşe TUZCU'nun 06/08/2019 tarihli ve 14531946 sayılı dilekçesi ile bildirilmektedir.

Uludağ Üniversitesi İngiliz Dili Eğitimi bölümü yüksek lisans öğrencisi Ayşe TUZCU'nun "Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi" konulu araştırmasını Yıldırım ilçesi Kaşgarlı Mahmud Anadolu Lisesinde uygulama yapma isteği ilimizde oluşturulan "Araştırma Değerlendirme Komisyonu" tarafından incelenerek değerlendirilmiştir. Araştırma ile ilgili çalışmanın **okul/kurumlardaki eğitim öğretim faaliyetleri aksatılmadan, araştırma formlarının aslı okul müdürlüklerince görülerek ve gönüllülük esası ile** okul müdürlüklerinin gözetim ve sorumluluğunda ilgi Genelge çerçevesinde uygulanması ayrıca **araştırma sonuçlarının Müdürlüğümüz ile paylaşılması** komisyonumuzca uygun görülmektedir.

Makamlarınızca da uygun görülmesi halinde olurlarınıza arz ederim.

OLUR
07.08.2019

Mustafa BİLİCİ
İl Milli Eğitim Müdür V.

Adres : Hocahasan Mh. İlkbahar Cad. No:38
(Yeni Hükümet Konağı A Blok) 16050/Osmangazi/BURSA
Telefon No:(0224) 445 16 00 Fax: 445 18 10
E-posta: arge16@meb.gov.tr İnternet Adresi: http://bursa.meb.gov.tr

Bilgi İçin : Engin SEYMEN
AR-GE VHKİ
(0224) 215 25 39

Appendix 16: Uludağ University Research Ethics Committee Approval Letter

BURSA ULUDAĞ ÜNİVERSİTESİ
ARAŞTIRMA VE YAYIN ETİK KURULLARI
(Sosyal ve Beşeri Bilimler Araştırma ve Yayın Etik Kurulu)
TOPLANTI TUTANAĞI

OTURUM TARİHİ	**OTURUM SAYISI**
31 Mayıs 2019	2019-04

KARAR NO 20 : Eğitim Bilimleri Enstitüsü Müdürlüğü'nden alınan Yabancı Diller Eğitimi Anabilim Dalı Yüksek Lisans öğrencisi Ayşe TUZCU'nun "Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi" konulu tez çalışması kapsamında uygulanacak ön test ve son test ürünleri yaratıcılık ölçme rubriği ve görüşme sorularının değerlendirilmesine geçildi.

Yapılan görüşmeler sonunda; Eğitim Bilimleri Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı Yüksek Lisans öğrencisi Ayşe TUZCU'nun *"Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi"* konulu tez çalışması kapsamında uygulanacak ön test ve son test ürünleri yaratıcılık ölçme rubriği ve görüşme sorularının, fikri, hukuki ve telif hakları bakımından metot ve ölçeğine ilişkin sorumluluğu başvurucuya ait olmak üzere uygun olduğuna oybirliği ile karar verildi.

Prof. Dr. Feridun YILMAZ
Kurul Başkanı

Prof. Dr. Abamüslim AKDEMİR
Üye

Prof. Dr. Doğan ŞENYÜZ
Üye

Prof. Dr. Ayşe OĞUZLAR
Üye

Katılmadı
Prof. Dr. Abdurrahman KURT
Üye

Prof. Gülay GÖĞÜŞ
Üye

Prof. Dr. Alev SINAR UĞURLU
Üye

Appendix 17: Institute of Educational Sciences Committee Approval

BURSA ULUDAĞ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

ENSTİTÜ YÖNETİM KURULU ARA KARARI

OTURUM TARİHİ	**OTURUM SAYISI**
05 Mart 2019	2019/09

KARAR NO: 02

Yabancı Diller Eğitimi Anabilim Dalı Başkanlığının 801610021 numaralı Yüksek Lisans öğrencisi Ayşe TUZCU'nun tez konusu önerisine ilişkin 26.02.2019 tarih ve 612 sayılı yazısı görüşmeye açıldı.

Yapılan görüşmeler sonunda Yabancı Diller Eğitimi Anabilim Dalı İngiliz Dili Eğitimi Bilim Dalı 801610021 numaralı Yüksek Lisans öğrencisi Ayşe TUZCU'nun "Yabancı Dil Olarak İngilizce Öğrenilen Sınıflarda Makine Çevirisinin Yazma Görevlerinde Yaratıcılığa Etkisi" isimli tez önerisinin BUÜ Lisansüstü Eğitim-Öğretim Yönetmeliğinin 28/1 maddesi uyarınca uygun olduğuna oy birliği ile karar verildi.

ASLI GİBİDİR
Raportör
Sadrettin ÖZDEMİR
Enstitü Sekreteri V.

Öz Geçmiş

Doğum Yeri ve Yılı: İzmir - 1987

Öğr. Gördüğü Kurumlar:	**Başlama Yılı**	**Bitirme Yılı**	**Kurum Adı**
Lise	2000	2004	Konak Anadolu Lisesi
Lisans	2004	2009	Trakya Üniversitesi
Ön Lisans	2006	2008	Anadolu Üniversitesi

Bildiği Yabancı Diller ve

Düzeyi : İngilizce – İleri

Çalıştığı Kurumlar	**: Başlama ve Ayrılma Tarihleri**	**Kurum Adı**
	1. 2009 – 2010	Ali Fevzi Ağan Ortaokulu
	2. 2010 – 2011	Alihocalar İlk/Ortaokulu
	3. 2011 – 2013	Vatan İlk/Ortaokulu
	4. 2013 – 2014	Hereke Anadolu Lisesi
	5. 2014 – 2017	Babasultan İlk/ortaokulu
	6. 2017 -	Kaşgarlı Mahmud Anadolu Lisesi

Yurt Dışı Görevleri:

Kullandığı Burslar:

Aldığı Ödüller:

Üye Olduğu Bilimsel ve Mesleki Topluluklar:

Editör veya Yayın Kurulu Üyeliği:

Yurt İçi ve Yurt Dışında Katıldığı Projeler: Comenius Hizmet İçi Eğitim

Comenius Okul Ortaklığı

Katıldığı Yurt İçi ve Yurt Dışı Bilimsel Toplantılar:

Yayımlanan Çalışmalar: Tuzcu, A. (2019). Foreign Language Education and its Cross-Curricular Links (Book Review). *Journal of Foreign Language Education and Technology*, *4*(2), 379-385.