



**THE REPUBLIC OF TÜRKİYE
MUĞLA SITKI KOÇMAN UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGE EDUCATION
ENGLISH LANGUAGE EDUCATION PROGRAM**

**CREATIVITY IN SECOND LANGUAGE SPEECH PRODUCTION:
INVESTIGATING THE ROLE OF TASK TYPE AND
PREPARATION TIME**

YUNUS EMRE KAHVECİ

MASTER THESIS

**JULY, 2024
MUGLA**

**THE REPUBLIC OF TÜRKİYE
MUĞLA SITKI KOÇMAN UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGE EDUCATION
ENGLISH LANGUAGE EDUCATION PROGRAM**

**CREATIVITY IN SECOND LANGUAGE SPEECH PRODUCTION:
INVESTIGATING THE ROLE OF TASK TYPE AND
PREPARATION TIME**

**YUNUS EMRE KAHVECI
ORCID ID: 0009-0001-4737-4312**

**THESIS SUPERVISOR
ASSOC. PROF. DR. PERIHAN KORKUT**

MASTER THESIS

**JULY, 2024
MUGLA**

JURY CONFIRMATION PAGE

This Master Thesis with the title of “Creativity in Second Language Speech Production: Investigating the Role of Task Type and Preparation Time” written by Yunus Emre Kahveci accepted as a Master Thesis in the Department of Foreign Language Education, and English Language Education Program by unanimous vote.

Prof. Dr. Turan PAKER
Department of Foreign Language Education
Chairman of the Jury

Assoc. Prof. Dr. Perihan KORKUT
Department of Foreign Language Education
Supervisor

Prof. Dr. Şevki KÖMÜR
Department of Foreign Language Education
Member

Date of defense: 16/07/2024

This thesis meets the necessary conditions of a Master Thesis in the English Language Teaching Program.

Prof. Dr. Şevki KÖMÜR
Director of Institute of Education Sciences

ETHICS STATEMENTS

This Master Thesis with the title of Creativity in Second Language Speech Production: Investigating the Role of Task Type and Preparation Time is declared to be written accordingly to the instructions of Muğla Sıtkı Koçman University Institute of Educational Sciences Thesis Writing Guide;

1. Data, information, and documents are gathered considering the ethical codes of conduct,
2. All of the information, documents, evaluations, and results are presented considering the ethical codes of conduct,
3. All of the resources used and benefited from are cited and referenced in the list of references,
4. No kind of change is carried out in the data used or the results emerged,
5. The work presented in this thesis is genuine,

And in the contrary case, all the results against the writer is accepted. 16 /07 /2024

Yunus Emre Kahveci

Use of chart, graph, photographs or any kind of content from this thesis without reference is protected by the law of Intellectual Property Rights no:5846.

ABSTRACT

CREATIVITY IN SECOND LANGUAGE SPEECH PRODUCTION: INVESTIGATING THE ROLE OF TASK TYPE AND PREPARATION TIME

YUNUS EMRE KAHVECI

Master Thesis, Department of English Language Education Program

Supervisor: ASSOC. PROF. DR. PERIHAN KORKUT

June 2024, 158 pages

The primary aim of this study is to investigate the impact of creativity, an individual difference often neglected in current research, on the second language (L2) speaking skills of EFL teacher candidates by comparing their speaking performance in a high-demand task and a low-demand task during the speech conceptualization processes. Additionally, this study examines whether the preparation time for tasks is related to the participants' creativity and speaking performance, and if so, which measures are associated. Addressing gaps in the literature, this study approaches the complex concept of creativity from both cognitive and personality dimensions, acknowledging that it cannot be observed through a single measure. A sample of 72 participants was selected using convenience sampling. To observe the cognitive dimensions of participants' creativity, the Alternate Uses Test (AUT) and Remote Associates Test (RAT) were employed, while the openness to experience subdimension of the Big Five Inventory-2 (BFI-2) was used for the personality dimension. For generating spoken data, an argumentative task (AT) was chosen as the high-demand task, and a picture narration (PN) task was chosen as the low-demand task. Participants were not given a time limit to prepare for the speaking tasks; their preparation times were recorded for analysis. A heavily quantitative approach was adopted for data analysis, converting participants' speaking data into numerical values based on complexity, accuracy, fluency (CAF), and discourse dimensions. Creativity skills were analyzed as predictor variables in terms of divergent thinking (DT) and convergent thinking (CT) for cognitive dimensions, and

openness for the personality dimension. Initially, correlation analysis was used to set out the relationships between predictor variables and CAF measurements, and significant correlations were noted for further analysis. Multiple hierarchical regression analyses (MRA) were employed to identify the extent to which significant correlations explained variance. Findings indicate that participants with a creative personality produced longer sentences in L2, articulated them more fluently, used more unique words, constructed sentences more accurately, and provided a greater amount of information in their speaking performance. However, negative correlations were found in terms of speech cohesion, suggesting that the connections between sentences were not sufficiently strong. Although this research is believed to be pioneering in the Turkish context, further studies are needed to expand the findings across different contexts and research settings to deepen the understanding of creativity's effects on L2 proficiency.

Keywords: creativity, speaking, speech conceptualization, CAF measures, foreign language teaching

ÖZET

YABANCI DİL KONUŞMA BECERİLERİNDE YARATICILIK: GÖREV TÜRÜ VE HAZIRLIK SÜRESİNİN ROLÜ AÇISINDAN BİR İNCELEME

YUNUS EMRE KAHVECİ

Yüksek Lisans Tezi, İngiliz Dili Eğitimi Bilim Dalı

Tez Danışmanı: DOÇ. DR. PERİHAN KORKUT

Temmuz 2024, 158 sayfa

Bu çalışmanın asıl amacı, günümüzde sıklıkla ihmal edilen bir bireysel fark olan yaratıcılığın, İngilizce öğretmeni adaylarının ikinci dil sözlü becerileri üzerindeki etkisini, konuşma kavramsallaştırma süreçlerinde yüksek bilişsel gereksinime ihtiyaç duyan bir konuşma görevi ve düşük bilişsel gereksinime ihtiyaç duyan bir konuşma görevindeki konuşma performanslarını karşılaştırarak araştırmaktır. Ek olarak, bu çalışma, görevlerden önceki hazırlık süresinin katılımcıların yaratıcılıkları ve konuşma performansları ile ilişkili olup olmadığını ve eğer ilişkiliyse, hangi ölçütlerle, ne derecede ilişkili olduğunu incelemektedir. Literatürdeki eksiklikleri tamamlamayı amaçlayan bu çalışma, karmaşık bir kavram olan yaratıcılığı hem bilişsel hem de kişilik boyutlarıyla ele almakta ve tek bir ölçütle gözlemlenemeyeceğinin bilinciyle gerçekleştirilmiştir. Elverişlilik örnekleme yöntemi kullanılarak 72 katılımcıdan oluşan bir örneklem seçilmiştir. Katılımcıların yaratıcılıklarının bilişsel boyutlarını gözlemlemek için Alternatif Kullanım Testi (AUT) ve Uzak Bağlantılar Testi (RAT) kullanılmış, kişilik boyutu için ise Beş Faktör Kişilik Envanteri-2'nin (BFI-2) deneyime açıklık alt boyutu kullanılmıştır. Konuşma verilerini toplamak için, yüksek bilişsel gereksinime ihtiyaç duyan konuşma görevi olarak açık uçlu tartışma görevi (AT) ve düşük bilişsel gereksinime ihtiyaç duyan resim anlatımı (PN) görevi seçilmiştir. Katılımcılara konuşma görevlerine hazırlanırken herhangi bir süre kısıtlaması getirilmemiş, hazırlık süreleri kaydedilerek analiz için saklanmıştır. Verilerin analizi için yoğun şekilde nicel bir yaklaşım benimsenerek, katılımcıların konuşma verileri karmaşıklık, doğruluk, akıcılık

(CAF) ve söylem boyutlarına göre sayısal verilere dönüştürülmüştür. Yaratıcılık becerileri bilişsel boyutlarıyla ıraksak düşünme (DT) ve yakınsak düşünme (CT), kişilik boyutuyla ise yeni deneyimlere açıklık olarak yapılan analizlerde yordayıcı değişkenler olarak kullanılmıştır. Verilerin analizinde başlangıçta korelasyon analizi kullanılarak yordayıcı değişkenler ve CAF verileri arasındaki ilişkiler belirlenmiş ve önemli korelasyonlar sonraki analizler için not edilmiştir. Korelasyon analizleri sonucunda aralarında önemli ilişkiler bulunan verilerin varyanslarının ne düzeyde açıklandığını belirlemek için çoklu hiyerarşik regresyon analizleri (MRA) kullanılmıştır. Bulgular, yaratıcı bir kişiliğe sahip olan katılımcıların ikinci dilde daha uzun cümleler kurduklarını, bu cümleleri daha akıcı bir şekilde kurduklarını, daha fazla nadir kelimeler kullandıklarını, cümleleri daha doğru bir şekilde kurduklarını ve konuşma performanslarında aktarılan bilgi miktarının daha fazla olduğunu göstermiştir. Öte yandan, konuşma bütünlüğü anlamında negatif korelasyonlar bulunmuş olup, cümlelerin birbirleriyle olan bağlantılarının yeterince güçlü kurulamadığı gözlemlenmiştir. Türkiye bağlamında öncü bir çalışma olduğu düşünülen bu araştırmaya rağmen, yaratıcılığın ikinci dil üzerindeki etkilerinin ileriki çalışmalarda daha fazla incelenerek farklı bağlamlarda ve araştırmalarda bulgularının genişletilmesi gerekmektedir.

Anahtar kelimeler: yaratıcılık, konuşma, konuşma kavramsallaştırma, karmaşıklık doğruluk akıcılık ölçütleri, yabancı dil öğretimi

ACKNOWLEDGEMENT

And now—Plato’s words mock me in the shadows on the ledge behind the flames: "...the men of the cave would say of him that up he went and down he came without his eyes..." (Keyes, 2004, p. 194)

I would like to begin by reminding that behind these lines, whenever read by myself or someone else, lie different people, worlds, and efforts behind a few hundred words, some numbers, and a few tables. And with the joy of leaving behind rigid objectivity, I want to add a few lines of my own to my work. I have reached these lines after many sleepless nights, much confusion, and perhaps a despair that, though not frequent, occasionally made its presence felt. It did not seem odd to me that a long letter of thanks, which should have been added at the beginning of the entire work, is being written at the end of everything. After all, this is a letter of thanks, and the many names that will appear in it will only be a few computer characters for the people reading these lines. However, each name mentioned here holds special significance for me and my story.

I want to begin to this lengthy acknowledgement by thanking my thesis advisor, mentor, and esteemed professor Assoc. Prof. Dr. Perihan KORKUT, who never hesitated to see me as one with herself. I owe a debt of gratitude to you and your family for doing everything you could to ensure the successful completion of this work, even in situations that were not your responsibility.

I also extend my thanks to Prof. Dr. Şevki KÖMÜR and Prof. Dr. Turan PAKER, who did not withhold their help, comments, and criticisms to ensure that my work was evaluated with an objective and professional eye.

I owe a debt of gratitude to every teacher who has touched my life throughout my educational journey, no matter how small. And has guided me, never withheld their help even for a moment, and taught and showed that people are much more than materials, classrooms, assessment criteria, and memorizing pages of writing. I especially thank Prof. Dr. Şevki KÖMÜR, Assoc. Prof. Dr. Perihan KORKUT, Assoc. Prof. Dr. Şeyda Selen ÇİMEN, Dr. Suat CAKOVA, Prof. Dr. Eda ÜSTÜNEL, Dr. Emel YILDIRIM, and Prof. Dr. Nevide AKPINAR DELLAL. I sincerely thank you once again for not only guiding

someone who wants to achieve their dreams but also for repeatedly reminding me why people cannot live without dreams.

I thank Muğla Sıtkı Koçman University for enabling my work and all my colleague candidates who voluntarily participated in my study. I also want to extend a special thanks to Dr. Mehmet ABİ, who often burdened himself with mathematical calculations, analyses, and statistics, for his help and the time he dedicated to teaching me how to interpret, analyze, and proofread the quantitative data in my study.

Perhaps the most special part of this gratitude section is thanking my family. I owe a debt of love, gratitude, and thanks to my mother, Şenay KAHVECİ, for illuminating my path whenever I would fall into darkness, for giving me the opportunity to follow my dreams, and for convincing me that no one can love another person as much as she loves me. I owe a debt of thanks to my father, Halit KAHVECİ, for teaching me how to be a righteous person and how to live an honorable and dignified life.

I owe a huge debt of gratitude to Serkan ŞAHİN, Muhammed Mirac GENÇ, Enes GENÇ, Oğuzhan KOÇ, Adnan TEKER, Eren YÜCEL, and many other friends whose names I cannot fit here, as well as my classmates with whom I shared an extraordinary but incomplete undergraduate life. They have been by my side for as long as I can remember, whom I consider family, who taught me the meaning of good and bad days, and with whom I learned what friendship truly means. I can rely on them blindly. I am complete with each of you and proud that our paths have crossed.

I need to extend my gratitude to a few special individuals in this section. I owe a special thanks to the teachers who showed me what teaching is not, those who prioritized their personal interests over professional ethics, made a habit of lying, and disgraced their profession. While the aforementioned individuals taught me what I should be, you achieved something even more valuable by showing me the kind of person I should never become, and for that, I am grateful to you.

Finally, I owe thanks to the Muğla Sıtkı Koçman University Scientific Research Project Office for their contributions to my master's thesis with the project number 23/164/14/4/[D6].

CONTENTS

	Page
JURY CONFIRMATION PAGE	iii
ETHICS STATEMENTS	iv
ABSTRACT.....	v
ÖZET	vii
ACKNOWLEDGEMENT	ix
CONTENTS.....	xi
LIST OF TABLES	xv
LIST OF FIGURES	xvii
LIST OF ABBREVIATIONS.....	xviii
LIST OF APPENDICES.....	xix
SECTION I	1
INTRODUCTION	1
1.1. Problem Status	2
1.2. Problem Statement.....	4
1.3. Purpose of the Research.....	5
1.4. Importance of the Research	6
1.5. Assumptions of the Research.....	8
1.6. Limitations of the Research	8
1.7. Definitions of the Research.....	9
1.8. Conclusion	10
SECTION II.....	11
LITERATURE REVIEW	11
2.1. Creativity	11

2.1.1. <i>Defining Creativity</i>	13
2.1.1.1. Two Aspects of Creativity: Divergent and Convergent Thinking.	16
2.1.1.2. A Social Personality Perspective: Openness	16
2.1.2. <i>Measuring Creativity</i>	17
2.1.2.1. Process	18
2.1.2.2. Person.....	19
2.1.2.3. Product	20
2.1.2.4. Environment.....	21
2.1.3. <i>Creativity and Language</i>	22
2.1.3.1. Creativity and Language Comprehension.....	23
2.1.3.2. Creativity and Language Production	28
2.2. <i>Speaking</i>	31
2.2.1. <i>Speaking as a Language Skill</i>	32
2.2.2. <i>Two In/Different Modes of Language Production</i>	35
2.2.3. <i>Measuring Speaking</i>	38
2.2.4. <i>Complexity, Accuracy, and Fluency</i>	42
2.2.4.1. Complexity.....	42
2.2.4.2. Accuracy	43
2.2.4.3. Fluency.....	44
2.3. <i>L2 Speaking and Creativity</i>	45
SECTION III.....	54
METHODS	54
3.1. <i>Research Design</i>	54
3.2. <i>Participants</i>	55
3.3. <i>Data Collection Tools</i>	56
3.3.1. <i>Personal Information Form</i>	56

3.3.2. <i>Remote Associates Test</i>	56
3.3.3. <i>Big Five Scale: Openness</i>	57
3.3.4. <i>Alternate Uses Test</i>	57
3.3.5. <i>Argumentative Speaking Task</i>	58
3.3.6. <i>Picture Narration Task</i>	59
3.4. <i>Data Collection Procedure</i>	59
3.5. <i>Data Analyses</i>	60
3.5.1. <i>Complexity</i>	61
3.5.1.1. <i>Syntactic Complexity</i>	61
3.5.1.2. <i>Lexical Complexity</i>	61
3.5.2. <i>Accuracy</i>	62
3.5.3. <i>Fluency</i>	63
3.5.4. <i>Discourse</i>	64
3.5.5. <i>The Analysis of Creativity Measures</i>	64
SECTION IV	66
FINDINGS	66
4.1. <i>Predictor Variables</i>	66
4.2. <i>The Interrelationship Among Predictor Variables</i>	67
4.3. <i>The Relationship Between Creativity and L2 Speaking Performance in AT</i>	68
4.4. <i>The Relationship Between Creativity and L2 Speaking Performance in PN</i>	75
4.5. <i>The Impact of Preparation Time on L2 Speaking Performance and Creativity</i>	83
SECTION V	85
DISCUSSION, CONCLUSION, and SUGGESTIONS	85
5.1. <i>Discussion</i>	85
5.1.1. <i>Discussion for the First Research Question</i>	85
5.1.2. <i>Discussion for the Second Research Question</i>	90

5.2. Conclusion	92
5.3. Suggestions	93
REFERENCES	95
APPENDICES	124



LIST OF TABLES

	Page
Table 1. <i>Descriptive Statistics of Placement Test Scores and Creativity Variables</i>	66
Table 2. <i>Correlations of Creativity Variables and Placement Test Scores</i>	67
Table 3. <i>Correlations of Creativity Variables and CAF Measurements in Argumentative Speaking Task</i>	68
Table 4. <i>Results of Hierarchical MRA Between Familiarity, L2 Proficiency and Openness in AT</i>	70
Table 5. <i>Results of ANOVA^a Between Familiarity, L2 Proficiency and Openness in AT</i>	70
Table 6. <i>Coefficients of the Hierarchical MRA Between Familiarity, L2 Proficiency and Openness in AT</i>	71
Table 7. <i>Results of Hierarchical MRA Between Total Number of Words Produced, L2 Proficiency, DT, and Openness in AT</i>	72
Table 8. <i>Results of ANOVA^a Between Total Number of Words Produced, L2 Proficiency, DT, and Openness in AT</i>	72
Table 9. <i>Coefficients of the Hierarchical MRA Between Total Number of Words Produced, L2 Proficiency, DT, and Openness in AT</i>	73
Table 10. <i>Results of Hierarchical MRA Between Causal Connectives, L2 Proficiency and DT in AT</i>	74
Table 11. <i>Results of ANOVA^a Between Causal Connectives, L2 Proficiency and DT in AT</i>	74
Table 12. <i>Coefficients of the Hierarchical MRA Between Causal Connectives, L2 Proficiency and DT in AT</i>	74
Table 13. <i>Correlations of Creativity Variables and CAF Measurements in PN</i>	76
Table 14. <i>Results of Hierarchical MRA Between Clause Length, L2 Proficiency and Openness in PN</i>	77

Table 15. <i>Results of ANOVA^a Between Clause Length, L2 Proficiency and Openness in PT</i>	78
Table 16. <i>Coefficients of the Hierarchical MRA Between Clause Length, L2 Proficiency and Openness in PT</i>	78
Table 17. <i>Results of Hierarchical MRA Between End-clause Pauses, L2 Proficiency, DT, and Openness in PN</i>	79
Table 18. <i>Results of ANOVA^a Between End-clause Pauses, L2 Proficiency, DT, and Openness in PN</i>	79
Table 19. <i>Coefficients of the Hierarchical MRA Between End-clause Pauses, L2 Proficiency, DT, and Openness in PN</i>	80
Table 20. <i>Results of Hierarchical MRA Between Disfluency, L2 Proficiency and Openness in PN</i>	81
Table 21. <i>Results of ANOVA^a Between Disfluency, L2 Proficiency and Openness in PN</i>	81
Table 22. <i>Coefficients of the Hierarchical MRA Between Disfluency, L2 Proficiency and Openness in PN</i>	82
Table 23. <i>The Descriptive Statistics of Students' Preparation Time</i>	83
Table 24. <i>Correlations of Creativity Variables and Preparation Time</i>	83
Table 25. <i>Correlations of CAF Measurements and Preparation Time</i>	83

LIST OF FIGURES

	Page
Figure 1. <i>Q-Q Plot for Placement Test Scores</i>	67



LIST OF ABBREVIATIONS

Abbreviations	Explanation
AT	Argumentative Task
AUT	Alternate Uses Test
BFI	Big Five Inventory
CAF	Complexity, Accuracy, Fluency
CT	Convergent Thinking
DT	Divergent Thinking
ELT	English Language Teaching
L1	First Language
L2	Second Language
MRA	Multiple Regression Analysis
PN	Picture Narration
RAT	Remote Associates Test

LIST OF APPENDICES

	Page
Appendix 1. <i>Ethics Committee Approval</i>	125
Appendix 2. <i>Personal Information Form</i>	126
Appendix 3. <i>Permission to Use RAT</i>	127
Appendix 4. <i>Sample RAT</i>	128
Appendix 5. <i>Sample RAT Answer Sheet</i>	129
Appendix 6. <i>BFI-2 Personality Inventory</i>	130
Appendix 7. <i>Sample AT Prompts (The Official Guide to the TOEFL Test)</i>	132
Appendix 8. <i>Sample PN Task (Adapted From Campbell, 2022)</i>	133
Appendix 9. <i>Sample Unpruned Transcription</i>	134
Appendix 10. <i>Sample AS-unit Transcription</i>	135
Appendix 11. <i>Sample Pruned Transcription</i>	136
Appendix 12. <i>Permission to Use Coh-Metrix Software</i>	137
Appendix 13. <i>Sample Weighted-clause Ratio</i>	138
Appendix 14. <i>Permission to Use Praat Software</i>	139

SECTION I

INTRODUCTION

Human history shows a persistent attempt to overcome obstacles posed by nature. Language has been essential to these attempts since it is a basic instrument for human survival and communication. Pinker (1994) asserts that communication skills are so fundamental that people frequently fail to recognize their complexity. According to Bailey (2004), people usually do not think about speech processes unless there's something strange about them. As a result, speaking is a fundamental part of daily existence.

This fundamental change naturally affected how language and its components are perceived, leading to a need for deeper analysis of different dimensions of speaking. According to Futrell et al. (2015), the primary objective of linguistics is to find explanations for the behavioural alterations in human languages, and one particular explanation is that language is functional in its nature; therefore, in his words: “language universals are grounded in the known properties of human information processing” (p. 10336). Furthermore, Lazaraton (2014) also makes an intriguing observation regarding the same issue:

When queried about competence in other languages, it would certainly sound odd to be asked, “How many languages can you understand (or read)?” makes more sense, but overwhelmingly the question put to us is “How many languages do you speak?” (Lazaraton, 2014, p. 106)

Speaking has been the most fundamental tool for human interaction. Similar to every foundation, speech has also different aspects. Brown and Yule (1983) put emphasis on the pragmatical side of speech, in terms of using it in information exchange. Goh (2016) also determines speaking as a highly complex skill which involves several

interactions of mental and social processes. In a parallel manner, Florez (1999) defines speaking as an interactive procedure of meaning construction whose form and meaning are determined by the context, participants, their experiences, the environment, and the intention of the speech. Although speaking is broadly defined as an innate mechanism that has coped with our evolution in time for interactional purposes, not every speech act is the same. For this reason, speech has been divided into several other subcategories by many researchers throughout time from transactional talk to interactional talk; talk as an interaction to talk as a transaction, and performance (Brown & Yule, 1983; Jones & Burnes, 1997; Richards, 2008). Due to its undeniable interlaced relation with mankind, speaking has drawn the attention of linguists and educational researchers for quite a considerable period. This peculiar bond maintained its significance further in educational settings-as in many other fields as well, leading to several methods and approaches in terms of teaching this unique skill.

This chapter provides a general overview of the study investigating the relationship between creativity and students' speaking skills, which will be explained and examined in greater detail within the following sections. Additionally, it discusses the research questions that form the foundation of the study, the rationale behind the need for such research, the limitations inherent in the study, and the general definitions of the terms used throughout the research.

1.1. Problem Status

Teaching speaking, along with other skills, has undergone various concepts of perception throughout time. As Pakula (2019) states, during the last 50 years, the conception of language proficiency has radically transitioned from focusing on the language as a linguistic system into focusing on the language for communicative purposes. In the meantime, the same shift has also been applied in teaching. By partially disregarding the linguistic aspect, Hinkel (2006) states that speaking is considered one of the four major language skills that are taken into account while learning and teaching a second language due to several factors such as the decrease in methods, the increase of the significance of bottom-up and top-down skills, recent perceptions of English, and the use of various skills integrated and taught throughout the context. According to Renandya

and Nguyen (2023), the primary objective of teaching speaking can be defined as producing several types of talk both accurately and precisely in a sociocultural context while doing it fluently. Thus, teaching speaking is considered to be one of the hardest skills to acquire, particularly for foreign language learners since they do not get to hear and use the target language regularly in naturalistic settings (p. 270). Rabab'ah (2005) also maintains the same idea by stating that the opportunity of using the target language is so limited that the only possibility for second language learners in Jordan to use their oral abilities in a natural interaction is to encounter native speakers who visit their country as tourists.

In addition to the lack of opportunities to use language in natural settings, several other factors affect the process of teaching speaking. Amoah and Yeboah (2021) divide those factors into two categories: linguistic factors (inadequate in using accurate and appropriate grammar, vocabulary, and pronunciation) and psychological factors (anxiety, unwillingness, shyness, and fear). Furthermore, Hinkel (2006) draws attention to the complexity of teaching speaking by highlighting the range and type of subskills that are involved in L2 oral production. Moreover, Gan (2012) draws attention to the fact that second language speaking is a highly complex skill due to its strenuous demands of both processing the task itself and processing another language which is quite flawed when the learner is involved in a speaking task. Additionally, Burns (2017) endorses this notion by stating in order to be a competent speaker of a second language, a learner must combine several skills, information, and procedures for the contexts of production. In addition, the result in speech must be “culturally and socially relevant, appropriate, and comprehensible to their interlocutors, as well as managing micro-level reactions and responses to what they utter” (p. 243). Furthermore, Bailey (2020) elaborates on the same subject by asserting that speech in a second language consists of “simultaneous deployment of linguistic systems” at different levels which include phonemes, morphemes, the lexicon, syntax, and discourse (p. 79). Therefore, speaking becomes a notably challenging skill to obtain.

Speaking is a challenging skill to master for a variety of reasons, but one prevalent aspect is individual differences among learners. While some learners speak with ease in front of large audiences, others might have problems even with basic tasks. As McDonough et al. (2015) states, individual factors affect miscellaneous aspects of learners' performance. These factors might be occurring due to various cognitive and

affective factors such as: the capacity of working memory, willingness to communicate, anxiety, and motivation (McDonough et al., 2015, p.188). Many learners feel more tense and anxious in foreign language classes compared to other classes, and this concern mainly stems from the speaking situation (Kitano, 2001, p. 549). A great number of studies were conducted to examine individual differences both from the linguistic and psychological perspectives (MacIntyre & Gardner, 1994; Dewaele & Furnham, 2000; Dörnyei & Kormos, 2000; Hedge, 2000; Rabab'ah, 2005; Elia & Irianti, 2015; Jiang & Dewaele, 2019; Amoah & Yeboah, 2021). However, creativity, which is one of the key aspects of individual differences, remains largely unexamined (Albert & Kormos, 2011).

1.2. Problem Statement

Otto (1998) states that, studies have been scarce even though the relation between creativity and success is undeniable in L2 learning. Research conducted on individual learner variables in second/foreign language learning has not fully explicated learners' creativity as a "potentially relevant variable" (Fernández-Fontecha, 2021, p. 279). According to Albert and Kormos (2011), the relevance of various cognitive, social, and psychological factors has been uncovered to account for derivative success in the second language; yet the importance of creativity has not been utterly explored, and neglected until the present day. Smith (2013) interprets the reason behind this information gap as the preconception of the inability to measure or assess creativity, and that it is seen as a soft construct by educators that does not fit for meticulous empirical studies.

Despite the inaccurate convictions, some scholars have highlighted the importance of creativity, especially in recent, communicative-based methods. Albert (2008) brought attention to the fact that creativity plays a crucial role in communicative learning environments which employ tasks that need students' imagination and originality. In a parallel manner, Sternberg (2002) implied that creativity was one of the major determiners of language acquisition. In these times, where teaching methods have transitioned towards communicative and student-centered approaches and tasks, the ability of student to express themselves in a language other than their own has become a significant attribute. These approaches and tasks frequently require students to retrieve or construct original ideas and therefore, it is rational to suggest that the results depend

greatly on learners' creative abilities. Learners' scarcity of creativity may conclude with difficulties while performing them, influencing negatively the development of L2 proficiency (Otto, 1998). In a corresponding way, an important articulation comes from Pishghadam and Mehr (2011):

In a nutshell, the notion that creativity has positive impact on learning and proficiency seems more acceptable as it can be claimed that language use is a creative act: we transform thoughts into language that can be heard or seen and likewise we are capable of producing sentences and even long texts that we have never heard or seen before. (Pishghadam and Mehr, 2011, p. 119)

Creativity is particularly evident in the cognitive processes. Studies (Bialystok et al., 2012; Kharkhurin, 2017) have shown that, bilingual or multilingual speakers' cognitive abilities such as mental control and cognitive flexibility are more extended compared to monolinguals, thus facilitating the creative processes in their mind. This suggests that creativity is an important difference among individual learners (Otto, 1998), and therefore, it is more crucial in second/foreign language settings. Even today, it is necessary to be aware of how little attention is actually dedicated to creativity, the sole factor that has brought humanity to this point, imparting the meaning and significance of being human. Albert (2012) also draws attention to the fact that in terms of empirical research, the relation between creativity and L2 has been "almost entirely neglected" and it is disappointing (p. 145). The narrow research and significance of creativity, a vital catalyst in the genesis of communication and consequently speech, both of which are integral to the essence of humanity, necessitates the undertaking of this academic study.

1.3. Purpose of the Research

The impact of creativity, as both a personality trait and a cognitive skill, on language use has been examined in numerous studies (Popov, 1992; Otto, 1998; Soveri et al., 2011; Wang, 2012; Ritchie et al., 2013; Mourgues et al., 2014; McDonough et al., 2015; Suzuki et al., 2022; Qin, 2024). The primary aim of this research is to investigate and analyse how participants' levels of creativity affect their speaking proficiency in their second language (L2) speech conceptualization processes by comparing spoken data from two different speaking tasks that require distinct levels of conceptualization. The

secondary aim of the research is to explore and analyse whether there is any relationship between the time participants spend preparing for these speaking tasks and their levels of creativity or the quality of their speaking performance. To investigate and reach the objectives of the study, the following research questions are addressed:

1. Which predictor variables of creativity demonstrate significant associations with CAF measurements while L2 proficiency is controlled?
2. In what ways does the task preparation time influence the quality of speech production and creativity?

1.4. Importance of the Research

The study has been designed with consideration of the limitations present in the existing literature on creativity and speaking skills. Therefore, it aims to address and fill the gaps identified in previous research in this field. Consequently, the study enhances its significance in the literature through its approach to addressing the identified problem, offering a more comprehensive understanding of the issue.

Previous research has predominantly viewed creativity as an outcome variable, often combined with other factors. One of the initial inquiries into creativity was conducted by Otto (1998). In this study, creativity was assessed with 34 Hungarian secondary class students. Utilizing a modified version of the Torrance Tests of Creative Thinking (TTCT), the researcher evaluated originality, problem-solving acumen, ideational fluency, and associative fluency, and then compared these scores with the students' grades. While a positive and reasonably strong connection emerged between these two variables, the study had inherent limitations due to being among the first to focus on creativity as a distinct aspect. A noteworthy omission in this study was the absence of tasks to assess language proficiency. Otto aimed to gauge language proficiency through scores from the Hungarian State Language Examination, which evaluated both oral and written performance. However, the method of scoring was not expounded upon. In contrast to Otto (1998), picture narrative and argumentative tasks will be employed in this study to gauge learners' second language spoken proficiency. The measurement of such data will be based on complexity, accuracy, fluency, and discourse aspects of the spoken data.

Another prominent study by Albert and Kormos (2004) delved into the connection between creativity and performance among 67 secondary school students using oral narrative tasks. However, this study found limited correlations between creativity scores and student performance. The researchers postulated that the relatively lengthy planning time allocated (5 minutes) could have influenced these outcomes. To avoid this type of limitation, no time restrictions were imposed on students in this study. Instead, the relationship between the time students used and their creativity, as well as their speaking performance, was investigated.

Smith (2013) also examined the relationship between creativity and L2 performance in open-ended tasks. The study involved 42 first-year university students and revealed a positive correlation, suggesting that students with higher creative abilities tend to be more proficient in L2. A notable limitation in this study pertains to the creativity tests employed. Typically, researchers refer to Guilford (1967) or Torrance (1966) for creativity tests; however, this study employed an alternative creative self-efficacy questionnaire for assessing creativity, with items translated by a proficient English user. Tests that pertain to personal psychology should ideally be translated and validated by experienced psychology researchers before use in academic research. Therefore, in contrast, all the creativity tests (Remote Associates Test, Alternate Uses Test, Big Five scale) used in this study are validated by experienced psychology researchers (Özen et al., 2015; Horzum et al., 2017).

Similarly, McDonough et al. (2015) explored the interplay between creativity and learners' performance in group problem-solving tasks. Their study involved 55 first-year Thai university students, and unlike Smith (2013), they employed the Torrance Tests of Creative Thinking to measure creativity. Results showed a positive relationship between creativity and learners' production of questions and coordination, but not with other language features identified as instrumental for task completion, such as "you" pronouns, subordinate reasoning clauses, conditionals, and stance. Since creativity is a dynamic process, the researchers acknowledged that relying on a single test to measure creativity might be inadequate. Hence, in this study, three diverse tests are incorporated to measure creativity from both convergent and divergent perspectives, thereby enhancing the validity of the results.

Recent research aligns with prior conclusions. For instance, Fernandez-Fontecha (2021) explored the connection between L2 semantic fluency, and creativity, with 35

Spanish EFL learners in the 12th grade. The study found that the most creative learners exhibited higher fluency in L2. One limitation was that proficiency was assessed based on written production and vocabulary size of the learners was not measured beforehand. Additionally, a significant individual difference, openness, which is one of the crucial measures of creativity, was not examined either. On the other hand, this study specifically focuses on openness and employs two open-ended oral tasks, offering better insight into conceptualization processes.

1.5. Assumptions of the Research

It is assumed that students' responses in group-administered creativity tests were not influenced by each other.

It is further assumed that students' speaking performance was not affected by external conditions. Specifically, it is assumed that anxiety and motivation did not prevent students from successfully completing the speaking tasks. The researcher is also assumed to maintain a genuine attitude, avoiding any behavior that could negatively impact students' speaking performance, and thoroughly addressing all questions to eliminate any doubts in the students' minds.

1.6. Limitations of the Research

The study has several limitations that may affect the generalizability of its results. Firstly, the research is conducted at Muğla Sıtkı Koçman University, a higher education institution in Türkiye, with first-year native Turkish speakers from the English Language Department. Therefore, further research is needed to determine how the findings might differ among participants with different native languages. Additionally, although it has been shown that the study's results are not affected by its effect size, the fact that it involves only 72 participants still impacts the generalizability of the findings. Consequently, studies with a larger number of participants are needed.

One of the creativity dimensions examined, divergent thinking, which is a cognitive creativity dimension, is assessed through an oral test in this study. Even though

divergent thinking can also be measured by figural tests, it was only examined in its verbal dimension in this study and therefore could not be investigated more comprehensively. This can be considered a limitation. Another creativity test to be completed by the students, the Remote Associates Test (RAT), consists of 120 items. It is assumed that students might become fatigued while completing this test since there is no validated and shortened Turkish version available in the literature. The analysis results showed that no significant outcomes were found for convergent thinking, as represented by the RAT. Consequently, the length of this test may have had a tiring effect on the students, and the inability to use a shortened version is another limitation.

Furthermore, creativity is taken as a predictor variable and is not examined as an outcome variable. Thus, since creativity should also be evaluated by the products it generates, the study's results apply only when interpreting creativity as a predictor variable.

In summary, while the study offers valuable insights into the effects of creativity on speaking skills, particularly in speech conceptualization processes, the aforementioned issues narrow the study's scope and consequently affect its generalizability.

1.7. Definitions of the Research

Although detailed definitions are presented in the next section, the frequently occurring terms within the research that need to be understood are briefly introduced under this heading.

Creativity. Despite the lack of a single agreed-upon definition, a frequently used definition exists. Plucker et al. (2004) define creativity as “the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context” (p. 90).

Divergent Thinking. Considered the fundamental cognitive dimension of creativity, divergent thinking is defined as the ability to generate multiple solutions to a given problem from various perspectives (Brophy, 2001).

Convergent Thinking. In contrast to divergent thinking, convergent thinking, another fundamental cognitive dimension of creativity, is the ability to develop the single

most appropriate solution to a problem from among many possible solutions (Brophy, 2001).

Alternate Uses Test. Initially proposed by Guilford (1959), this test is used to measure divergent thinking skills. Participants are asked to continuously list how a given object can be used other than its intended purpose within a certain time frame.

Remote Associates Test. Developed by Mednick (1962), this test aims to measure convergent thinking skills. Participants are expected to find another word that completes three given words meaningfully.

Big Five Inventory-2. A personality inventory further shortened by Soto et al. (2017) from the original BFI personality inventory. Only the openness to experience dimension of the personality inventory is used in this study.

L1. Used to define an individual's native language.

L2. Defines any language other than an individual's native language.

The definitions provided under this heading do not go beyond outlining a general framework. The actual definitions of the terms, how they originated, and their roles in the study are detailed in the next section.

1.8. Conclusion

In the introduction part of the study, discussion revolves around the significance of language, the primary attribute humanity has discovered/produced since its inception, and its role in creating meaning through the usage of language, the communication that arises from making associations among these meanings, and how speech, as one of the four fundamental functions of language, is incorporated into educational processes. It further delves into how the individual differences of students within these educational processes influence the processes, why addressing the concept of creativity as an individual difference is important, and what role creativity plays in second language education. This overarching framework serves to outline the requirements of this study.

SECTION II

LITERATURE REVIEW

This section provides explanations regarding the purpose and definitions underlying the discussed concepts that will enhance the clarity of the study, thus articulating its scientific contributions more explicitly. In addition, it will summarize the relevant studies in the literature.

2.1. Creativity

While initially, every definition regarding creativity may be accepted to some extent, establishing an absolute definition of creativity is a challenging task. One might argue that creativity is an important personality trait, a universal human quality, and it distinguishes individuals as more “special”. However, despite drawing a very broad framework, considering the historical context of the concept of creativity, it becomes apparent that these qualities we commonly acknowledge as norms are actually exceptional qualities and that creativity is a merit that has been discovered quite late (Glaveanu & Kaufman, 2019). Similar to its definition, the history of creativity, is diverse and complex. Before making a definition in the modern sense, an examination of the role that creativity has played throughout the human history is necessary. Despite diverse narratives about the place of creativity in human history, there exists a commonly used primary narrative (Dacey, 1999).

Traces of creativity have been evident since the earliest moments of humanity. Even the most primordial ancestors of modern humans, often perceived as primitive and resistant to change, exhibited creativity when creating alphabets, transitioning to settled societies, and organizing their armies systematically, shaping creativity both in concrete

and abstract terms. However, an important question arises here: What distinguishes Michelangelo, or Bach, or Ibn Sina from other individuals? Why are these individuals associated with creativity more while others are marginalized, even differentiated from each other?

As previously stated, creativity, a trait that modern humans deem normative and prioritize as a necessity in both personal and professional life, was historically perceived as a divisive characteristic among individuals until recent times. The gift of fire stolen by Prometheus from the gods opened the doors to science and art for humanity, yet Prometheus was cursed for this altruistic act. Creativity was viewed as a divine attribute, and being overly creative was believed to disrupt the balance of the universe, considered as a passionate disaster to be avoided (Mason, 2003). These beliefs, passed down through generations until the Middle Ages, were reinforced by the pressure exerted by churches, despite being the places where many masterpieces in art forms such as architecture, sculpture, and painting have been produced and acclaimed as masterpieces in modern times, leading to the belief that creativity was entirely a reflection of God's presence in mankind (Kaufman & Glaveanu, 2019).

Even during the Renaissance period, creativity was approached with scepticism, but at least it was now perceived not as a reflection of God but as a gift from Him. The idea of individualism, which gained importance with the Industrial Revolution, brought about significant technological and social issues, emphasizing problem-solving skills and consequently highlighting creativity as a pivotal talent (Dacey, 1999). As time progressed, accessing different cultures and lifestyles became easier for various people. Access to different cultures provided a conducive infrastructure for creating and enhancing a more universal, shared culture. Today, every small action taken to develop and disseminate this shared culture such as sharing a photo online or commenting on a post can be regarded as a micro-example of creativity (Glaveanu & Kaufman, 2019).

As previously mentioned, creativity has been subject to a variety of divisions throughout its history, including the question of whether it is a social phenomenon or an individual phenomenon, whether the products of creativity should be considered innovative or valuable, and whether creative actions should be implemented or remain mere ideas (Glaveanu, 2013; 2016). Even today, especially due to interdisciplinary approaches to defining creativity, it is categorized into different dimensions such as purely cognitive and social personality dimensions (Sawyer, 2012).

2.1.1. Defining Creativity

It has been noted that the definition of creativity is as diverse and complex as its background. Examining and interpreting creativity as a scientific element was not achieved until modern times. After the decline of behaviorism's influence, creativity began to be extensively studied both in more specific domains (Guilford, 1950; Renzulli, 1994; Torrance, 1962) and in broader contexts (Blicbau & Steiner, 1998; Bloland, 1987; Innamorato, 1998). While many comprehensive studies were expanding the place of creativity in the scientific literature, they also brought along many challenges for this new concept. Plucker (2004) also seems to have noticed this, as evidenced by the first significant study to articulate how crucial creativity is for an individual's problem-solving skills, emotional and social state, and academic and social abilities, yet also pointing out the limitations of studies in this field due to their perspective on creativity.

In their study, Plucker et al. (2004) suggest that, while research on creativity adequately observes the potential forms of its application, it argues that this potential is not carried out thoroughly. It is indicated that the underlying issue lies in the perspective towards creativity, stemming from both researchers and laypeople. Furthermore, it is contended that this perspective diminishes researchers' enthusiasm for studying creativity. Additionally, this problem is examined under four main headings:

1. *Myth A.* Individuals are either born creative or lacking in creativity,
2. *Myth B.* Creativity and creative individuals are characterized by negative attributes both in the field of psychology and within society,
3. *Myth C.* "Creativity is a fuzzy, soft construct" (Plucker et al., 2014, p.86)
4. *Myth D.* Creativity is more productive within a group, enhancing its impact within the group rather than individual creativity.

Attempting to demonstrate that these four myths are not as researchers suggest, Plucker et al. (2004) show that the absence of a standardized definition of creativity lies at the core of this notion. To substantiate this idea, a two-stage content analysis was planned. In the first stage, a sample of 30 articles (10 from each of the fields of psychology, education sciences, and business) published between 1996 and 2002 containing the term "creativity" in their titles were selected to calculate the frequency of the term "creativity" appearing in the articles, whether any implicit or explicit definitions of creativity were present, and how often the term "creativity" was used before a

definition was given. In the second stage, the same procedures were repeated with an additional sample of 60 articles published in two main journals on creativity (for the year 2004, *Creativity Research Journal* and *Journal of Creative Behavior*). According to the results of the study, out of the 90 articles, only 34 (38%) provided an explicit definition of creativity, 37 articles (41%) implicitly provided a definition, and 19 articles (21%) did not contain any definition. Upon further analysis, it was observed that each study explicitly defining creativity mentioned common themes (uniqueness and usefulness) but did not offer similar definitions. With these data, Plucker et al. (2004) argue that clear and shared definition of creativity is lacking in studies related to creativity, thus suggesting that under the umbrella of creativity, vastly different topics are discussed, and creativity is approached from various perspectives. In summary, if metaphorically encapsulated within the study:

“This is not merely a case of comparing apples and oranges: We believe that this lack of focus is tantamount to comparing apples, oranges, onions, and asparagus and calling them all fruit. Even if you describe the onion very well, it is still not a fruit, and your description has little bearing on our efforts to describe the apple.” (Plucker et al., 2004, p. 88-89)

In the study where the debate on the necessity of a common definition of creativity was initiated, Plucker et al. (2004) proceeded by identifying commonalities in the examined articles to propose a comprehensive definition of creativity: “Creativity is the interaction among aptitude, process, and environment by which an individual or group produces a perceptible product that is both novel and useful as defined within a social context.” (Plucker et al., 2004, p. 90)

In addition to this definition, it is emphasized that in future studies focusing on creativity, researchers should explicitly define creativity, refrain from evaluating participants’ creativity solely based on the results of creativity tests, specify the similarities and differences of their definitions with those made by other researchers, and particularly highlight for whom and in what context creativity is being used.

Plucker et al. (2004) aimed to establish guidelines for future studies focusing on creativity with their study. However, six years later, as indicated by Hennessey and Amabile (2010) in a similar study by Simonton (2016) in another study twelve years later,

consensus on criteria and definitions has still not been reached. A more comprehensive study conducted recently further substantiates these findings.

In a study aimed at providing a broader portrait of Plucker et al.'s (2004) work, Puryear and Lamb (2020) conducted an analysis method parallel to the content analysis used by Plucker (2004), examining 600 articles. According to the analysis results, 53% of the 600 studies provided an explicit definition of creativity, while it was noted that there was a noticeable decrease in the number of studies that did not provide any definition compared to the data in Plucker's study. Puryear and Lamb (2020) mention that each of the studies they examined opted for a specific definition of creativity tailored to their respective fields, indicating that they approached creativity from different perspectives specific to these fields, and add: "...while it is entirely plausible that different contexts require different conceptions, whenever researchers stray from the core, it necessitates that they be more mindful that there are potential content validity ramifications in doing so." (Puryear & Lamb, 2020, p. 8)

Based on the results of the aforementioned studies, it is understood that a common definition of creativity may not be applicable to every study content. In this study, as it is suggested by various studies on creativity (Simonton, 2012; Kumar & Kumari, 2016; Javaid & Pandarakalam, 2021; deVink et al., 2022), it is principally aimed to evaluate participants' creativity more comprehensively and integratively by utilizing two diverse perspectives of cognitive creativity. These perspectives are divergent thinking, long regarded as the main construct of creative thinking, along with convergent thinking, referred to as a contrasting thinking style to divergent thinking (Guilford, 1950). In order to fully evaluate an individual's creativity, personality traits should not be overlooked alongside the cognitive dimensions (Simonton, 2012). As observed in McRae's (1987) study, the personality trait with which the cognitive dimensions of creativity most vigorously correlate is "openness,"¹ which measures an individual's receptiveness to new experiences. Therefore, in order to better align with the purpose of the research and provide a more explanatory approach, rather than defining creativity under a unitary definition within this study, it is proposed to address various dimensions of creativity, each to be defined within its own framework. Thus, taking into account previous researchers' suggestions, participants' creativity aptitudes will be approached within a broad framework without being confined to a single criterion.

2.1.1.1. Two Aspects of Creativity: Divergent and Convergent Thinking. The terms coined by Guilford (1959), namely divergent thinking (DT) and convergent thinking (CT), are regarded as two distinct dimensions of creativity. Divergent thinking, is the process of generating original and varied solutions for a given problem. Convergent thinking, on the contrary, is the process of selecting the most appropriate solution (or the few most appropriate options) among the solutions proposed for the given problem (Brophy, 2001). Although these two concepts, by definition, seem to converge at a common point, researchers have often contrasted them. Despite the fundamental notion in idea generation that more ideas are always better, this alone does not necessarily constitute a creative process (Qian et al., 2019). Kaufman et al. (2010) also discuss the intriguing yet insurmountable irony arising from this situation in scientific research on creativity. While these two concepts are often distinguished in cognitive theories (Campbell, 1960; Smith, 1998; Fasko, 2001; Scott et al., 2004; Simonton, 2007; Weisberg & Hass, 2007; Wolf & Mieg, 2010; Van de Kamp et al., 2015), they are considered inseparable dualities in Guilford's (1967) triangular theory. De Vink et al. (2022) also emphasize the need to consider both types of thinking together, arguing that, according to the definition of creativity based on Plucker et al.'s (2004) definition, it requires not only novelty but also practicality. In a parallel vein, Javaid and Pandarakalam (2021) have discussed that "genuine creativity" should encompass both types of thinking, stating that creativity is not a dilemma between divergent thinking and convergent thinking but a "*harmonious blend*" between these two ideas (p. 138). For these reasons, in the present study, participants' levels of creativity were measured using Guilford's (1959) proposed divergent thinking and convergent thinking as cognitive criteria.

2.1.1.2. A Social Personality Perspective: Openness. Although the impact of cognitive differences on creativity has been frequently investigated, the influence of individuals' personality traits on their creativity is undeniable. Studies (McCrae, 1987; Silvia et al., 2009; Chamorro-Premuzic & Furnham, 2009; Beaty et al., 2014; Kaufman, 2016) have consistently shown a strong positive correlation between "openness to experience" and an individual's cognitive creativity. To provide additional elucidation, openness is one of the domain scales of a test commonly used in psychology assessments for character analysis called the "Big Five Inventory," which examines individual differences (Dewaele, 2012). This fundamental trait of "Open-Mindedness," also known as "Openness" or "Openness to Experience" (Costa & McCrae, 2008), reflects an

individual's constant quest for new experiences and curiosity about the unknown. Individuals with high openness scores are often described as more curious, reflective, and imaginative, while those with lower scores are described as less curious and less inclined towards art and analytical pursuits (Dewaele, 2012). Additionally, Runco and Acar (2019) note a logical connection between DT and openness, suggesting that individuals open to diverse experiences are also open to various ideas, making it easier for them to develop such thoughts. Furthermore, Ivcevic and Hoffman (2019) highlight openness as the most consistent and robust personality trait that aligns with cognitive criteria for creativity. For these reasons, in this academic study, openness is preferred as a personality criterion alongside cognitive criteria when classifying participants' creativity.

2.1.2. Measuring Creativity

As previously mentioned, not only is there no consensus among scientific communities on a single definition or historical context for creativity, but there are also disparate approaches and tools for its measurement. The misconception that creativity is too complex to be measured, and the resulting ambiguity, primarily stem from the lack of a unified definition among researchers (Batey, 2012; Piffer, 2012). It is widely accepted that the formal inclusion of creativity as a distinct field of scientific inquiry began with Guilford's 1950 presidential address to the American Psychological Association (APA). Therefore, categorizing the intensity of scientific research on creativity as pre- and post-Guilford is not unreasonable. However, Taylor and Barron (1963) argue that Galton's 1883 publication, *Inquiries into Human Faculty*, already posed significant questions about the measurement of creativity. Additionally, Torrance (1982) noted that as early as the 1930s, Guy M. Whipple was developing various tests focused on imagination and creativity. While creativity has always been a subject of scientific interest, its emergence as a primary research focus is relatively recent. Researchers such as Amabile (1983) and others (Rhodes, 1961; Couger et al., 1993; Sternberg & Lubart, 1995; Fishkin & Johnson, 1998; Thompson & Lordan, 1999; Horn & Salvendy, 2006; Barbot et al., 2011; Glăveanu, 2013) have advanced creativity metrics beyond cognitive and personality approaches, developing more integrative sociocultural theories that encompass creative thinking. Furthermore, according to Sternberg (2010, 2018), both creativity and its assessments are currently experiencing an unprecedented period of activity.

Criteria for measuring creativity are examined under four main categories, paralleling the various areas in which the definition of creativity is distributed:

1. Process,
2. Person,
3. Product,
4. Environment (Metwaly et al., 2017; Plucker et al., 2019).

2.1.2.1. Process. In the process approach used by researchers, the primary aim is to observe cognitive processes during creative production by presenting participants with problems that encourage the generation of as many responses as possible to a stimulus. This approach, based on DT, contrasts with other standardized tests by requiring participants to produce multiple answers. Since the early periods of creativity research, DT has been a predominant focus and remains a significant criterion for measuring creativity (Kaufman et al., 2008). Prominent DT tests include the Structure of the Intellect (Guilford, 1967), Wallach-Kogan Creativity Tests (Wallach & Kogan, 1965), the Torrance Tests of Creative Thinking (Torrance, 1966, 1974, 2008), Creativity Assessment Packet (Williams, 1980), and Scientific Creativity Test (Hu & Adey, 2002). Although the application methods and instructions of these tests differ, their evaluation criteria remain consistent. The results of DT tests are analysed across four dimensions: fluency (the number of responses to a stimulus), originality (the statistical uniqueness of the responses), flexibility (the number of different unique categories the responses can be classified into), and elaboration (the detail of the responses) (Metwaly et al., 2017; Plucker et al., 2019). Therefore, both the quantitative and qualitative aspects of the responses are significant. As Plucker et al. (2019) noted:

“For example, if a person were trying to decide what to buy as a birthday present for her brother, she could come up with as many ideas for presents as possible (fluency), presents that no one else would think of (originality), a list of different types of presents he may like (flexibility), or a list of the different basketball-related presents he might like (elaboration). However, in this example, as in life, choices have to be made eventually, and evaluative (convergent) thinking must be done to select the actual gift to be purchased.” (Plucker et al., 2019, p. 46)

Therefore, the creativity levels of participants observed solely through DT cannot be fully measured. According to Metwaly et al. (2017), numerous studies using the same

tests have yielded contradictory results. Particularly concerning discriminant validity, originality and fluency scores have led to problematic outcomes in many studies. For instance, Silvia et al. (2008) found a high correlation between originality and fluency scores, raising questions among researchers about whether fluency explains anything beyond its own dimension. Additionally, as seen in various studies, when fluency scores are controlled for or when the number of participants increases, the reliability of originality subscores diminish. Furthermore, according to Zeng et al. (2011), although there are four stages in the creative thinking process (problem analysis, ideation, evaluation, and implementation), DT tests focus solely on the ideation stage, neglecting the other stages. Recent studies (Primi, 2014; Silvia, 2015) suggest that DT might be more closely related to intelligence than previously thought, implying that variability in creative ability might need to be examined through intelligence rather than DT. In light of all this information, it is evident that while seeking a correlation between DT and creative abilities, other factors should not be overlooked.

2.1.2.2. Person. In the person approach, unlike the process approach, the focus is on the participants' characteristic traits, evaluating their creativity through self-report questionnaires or examining past behaviours (Metwaly et al., 2017; Plucker et al., 2019). The criteria are typically based on common behaviours exhibited by creative individuals, assuming that participants who exhibit similar behaviours can be identified as creative. Examples of such criteria include the Creative Personality Scale (Gough, 1979; Kaufman & Baer, 2004), Self Report of Creative Traits (Runco et al., 2017), and the Adjective Check List (Gough, 1979; Domino, 1994). However, according to Metwaly et al. (2017), merely possessing these traits does not necessarily indicate active engagement in creative thinking processes. Additionally, even if individuals possess these traits, they may lack the necessary environment or motivation to engage in creative activities. Thus, merely possessing the characteristics of creative individuals does not make them inherently creative (p. 251). Priest's (2006) study supports this notion. In a study comparing self-assessments of musical compositions by 54 undergraduate students with expert evaluations, significant statistical differences were observed between the classification of the students and the expert evaluations. Similar significant differences were found in studies by Reiter-Palmon et al. (2012) and Kaufman et al. (2010). Consequently, such results raise questions about the validity of conclusions drawn from these studies. Furthermore, as with all self-reports, there is a potential risk of bias in those related to

creativity (Kaufman et al., 2007), meaning that participants, either knowingly or unknowingly, may distort the data for various reasons.

Self-assessment tests have been shown to produce results that are not entirely valid for determining individuals' creativity. However, the Big Five Inventory (BFI) (John et al., 1991) is a widely accepted personality analysis test in psychology, and because it includes a subdimension of openness to experience, it is also considered applicable in the field of creativity. In a meta-analysis conducted by Karwowski and Lebeda (2015), a strong correlation was found between the personality traits in the BFI and domain-general measures of self-assessed creativity. Based on the findings of this study and others (Silvia et al., 2009; Beaty et al., 2014; Kaufman, 2016), the openness to experience dimension of the BFI is considered a valid criterion for creativity tests.

2.1.2.3. Product. Creative thinking processes cannot be considered or evaluated in isolation from their outcomes (Horn & Salvendy, 2006). Therefore, to assess creativity, a product-based approach is recommended by many researchers (Metwaly et al., 2017), although its importance is often insufficiently emphasized (Plucker et al., 2019). Despite the development of advanced and contemporary assessment products (such as the Creative Product Semantic Scale, Besemer & O'Quin, 1999; Student Product Assessment Form, Reis & Renzulli, 1991), the most frequently used scale for product-based approaches is the Consensual Assessment Technique (CAT) developed by Amabile (1982). The primary objective of this scale is to eliminate the ambiguity surrounding what is being measured by having appropriate observers independently evaluate the resulting product to determine its creativity. Essentially, CAT replicates real life to some extent, as people assess the usefulness, creativity, and artistic personality of others based on the products they create. In CAT tests, participants are expected to produce a product which can be a story, poem, drawing, or similar, and experts in the field then evaluate the creativity of the product using a Likert scale (Barbot et al., 2011). The products are evaluated by comparing them to each other without a fixed standard, and the results are determined by examining the inter-rater reliability scores. Once sufficient reliability is achieved, the final assessment of the product is made by averaging the given scores (Amabile, 1982). As Plucker et al. (2019) noted, this technique has been successfully applied in various contexts and different research settings. Despite its successful applications, some concerns about CAT have been raised by researchers.

Although Amabile (1996) emphasized the necessity of specific formal training for individuals evaluating creative products (p. 73), she did not specify the extent of this training. Kaufman et al. (2008) demonstrated in their study that when the creative product being evaluated was a poem, the evaluation criteria of experts and novices were unrelated, whereas, for a short story, the evaluations showed a high correlation. Similarly, in a study involving three different groups (professional film critics, a movie website reviewers, and college students) assessing films, Plucker et al. (2009) observed similar inconsistencies. Kaufman and Baer (2012) also noted that the required level of expertise varies by field, and experts assessing CAT tests should be evaluated according to the specific domain. Tan et al. (2015) observed that the personality traits of evaluators significantly influence their assessments. Evaluators with high agreeableness levels, as determined by BFI tests, were found to be more lenient in their evaluations. Additionally, cultural differences among evaluators (Niu & Sternberg, 2001), the extent to which evaluators accept the ideas being assessed (Runco & Smith, 1992), and subjective perspectives on evaluation criteria (Lindauer & Long, 1986) are some of the factors that hinder reaching a consensus among experts.

2.1.2.4. Environment. The environment approach examines the context in which an individual's, or group's creativity is recognized, and thus, its criteria establish an indirect rather than a direct relationship between environmental factors and creativity (Batey, 2012). In studies employing the environment approach, researchers investigate how individual, organizational, or cultural influences affect creativity (Metwaly et al., 2017). In a study aiming to identify factors fostering creativity in educational settings, Davies et al. (2013) linked these factors to an appropriate physical environment, the effective use of this physical space, the sufficiency of materials in this environment, an engaging and enjoyable pedagogical approach, and the optimal use of time to enhance student-teacher relationships. Several different measures have been developed for the environment approach (e.g., the Work Environment Scale, Moos, 1986; the Work Environment Inventory, Amabile & Griyskiewicz, 1989; the Team Climate Inventory, Anderson & West, 1998; the Situational Outlook Questionnaire, Isaksen et al., 1999), but the most commonly used measure is Assessing the Climate for Creativity (KEYS) by Conti et al. (1996). These measures generally aim to gather individuals' perceptions of resources, leadership, freedom, and creativity in the work environment. According to Hoelcher and Schubert (2015), most approaches to measuring creativity focus extensively

on the intrinsic dimensions of an individual's creativity such as cognitive or character traits, neglecting environmental factors. Consequently, more research is needed to observe the effectiveness of the environment approach. Mathisen and Einarsen (2004) argue that there is a controversy regarding whether the entire organization or just a specific group should be examined when considering environmental factors. Another critical point highlighted by Metwaly et al. (2017) is that evaluations might vary as individuals interpret their sociocultural and economic situations differently. Therefore, the creativity levels of individuals in the same work environment might reflect their internal states rather than the work environment itself.

In conclusion, while research on creativity has visibly increased in recent years, it has not shown significant methodological progress. The absence of a universally accepted definition of creativity in the literature, the reliance on domain-specific rather than multidimensional measures, and ongoing debates about the validity of these measures have hindered substantial advancement beyond the point left by Plucker et al. (2004). Consequently, this academic study aims to assess participants' creativity as a "holistic but multi-faceted construct" (Suzuki et al., 2022, pp. 1323) by considering three different dimensions (DT, CT, and openness) rather than focusing on a single aspect.

2.1.3. Creativity and Language

From the early sections of the study, it has been discussed that there is no uniform, universally accepted definition of creativity. Since Guilford highlighted creativity as a serious research field in the 1950s, studies have predominantly focused on the DT aspect of creativity. However, due to the current research, creative thinking is a far more complex concept, surpassing the DT dimension, is a reality acknowledged by the scientific community (Silvia, 2015). Recent studies have demonstrated that achieving a creative outcome depends on the ability to employ numerous cognitive processes, including DT, CT, flexible thinking, set shifting (i.e., cognitive flexibility), and flat and broad associative hierarchies simultaneously (Sio & Ormerod, 2009; Baas et al., 2015). Therefore, it can be stated that an individual's creativity, and thus creative thinking, requires the coordinated use of various executive functions within cognitive processes (Cushen & Wiley, 2011). As Vartanian (2009) indicated, flexibility is a crucial concept for creative processes. Zabelina and Robinson (2010) also emphasized the importance of

flexibility, stating that creative individuals exhibit more flexible cognitive control. By managing their executive functions, they transform the process from an automatic operation into a highly associative, sustainable, and goal-oriented procedure.

Considering language as a mode of thinking, interesting results emerge regarding flexibility. According to Takeuchi et al. (2011), creative processes require a well-developed working memory. Baddeley (2003) also noted that working memory is a critical system for higher-order cognitive processes such as reasoning, problem-solving, and comprehension. Thus, working memory serves a crucial function in creative processes, which necessitate the simultaneous processing of various functions. Studies have shown that working memory is significantly more developed in bilingual individuals compared to monolingual individuals (Morales et al., 2013; Birke Hansen et al., 2016). Furthermore, Carlson and Meltzoff (2008) found that early bilingual individuals possess a more developed working memory than later bilingual individuals. Numerous studies have observed that bilingual individuals use flexibility more effectively than monolingual individuals because they constantly switch between two languages (or more, if plurilingual) (Soveri et al., 2011; Verreyt et al., 2016). Additionally, Yang et al. (2021) observed that as proficiency in a second language increases, so do an individual's creative thinking skills. Consequently, these studies indicate that bilingual individuals are more adept at the cognitive processes required for creative thinking and at transitioning between these processes.

The primary and most important purpose of using a language is communication. And communication necessitates a bidirectional interconnection and intercondition. During this interconnection and intercondition process, not only does information transfer occur, but various psychological and cognitive procedures also play significant roles. According to acquisition theory, linguistic comprehension precedes linguistic production because effective expression cannot occur without sufficient comprehension (Nan, 2018). Therefore, when examining the impact of creativity on language, it is imperative to consider its effects on both comprehension (reading, listening) and production (writing, speaking) aspects of language.

2.1.3.1. Creativity and Language Comprehension. Listening and reading skills involve similar psychological activities and decoding processes. Despite the conventional wisdom that listening and reading are receptive skills, individuals do not merely respond passively to stimuli. Instead, they actively engage in the process by combining external

stimuli with their predictions, experiences, and schematic knowledge. When individuals are listening, they must comprehend the incoming information in a very short time, as they usually do not have the opportunity to slow down, pause, or replay the material. This rapid and complex mode of thinking not only enhances the listener's listening skills but also allows the individual to read more quickly by increasing their capacity to respond to language. On the other hand, extensive reading exposes individuals to schematic knowledge continuously, making the process of accessing to prior knowledge more automatic. Therefore, when they are listeners and acquire new information, their ability to associate this information with their previous knowledge improves. Additionally, since the micro-skills used in reading and listening processes are the same, individuals can quickly identify key words, make references, and predict while both reading and listening (Nan, 2018).

Most empirical research on creativity and the four language skills has focused on language production rather than comprehension. One significant reason for this is that creativity, by nature, is more observable through output than input, making it a more apparent personality trait or/and cognitive process. However, despite being largely overlooked, there are still studies on the relationship between input and creativity. Popov (1992) conducted one of the first significant studies, investigating the connection between students' reading skills and their creativity. The researcher initially observed the translation styles of students. Where translation style referred to the personal approach of the students regarding to translation rather than their translation styles. After that, the researcher categorized them into four groups: objectors (students bored with translating, attending classes only to complete translations unwillingly to pass the class), nonsense producers (students who, despite their efforts, produced nonsensical translations), sense producers (students who produced generally meaningful translations with minor errors), and guessers (students who could guess from context, understand the text quickly, and summarize it concisely). Based on these observations, Popov decided to investigate whether there was a relationship between reading skills and creativity among 63 freshman engineering students. For reading skills, measures such as overall text comprehension, predicting the topic from the title, predicting the topic from key words, sentence completion, paragraph completion, segmenting the text into paragraphs, writing abstracts, finding the main idea, and writing a critical response to the text were selected. For creativity scores, Guilford's creativity test was chosen. The findings indicated a high

correlation between the skills of writing abstracts and critical responses and students' creativity scores. In the critical writing task, the researcher emphasized that conflicting texts were provided, requiring students to notice details and write their critiques based on these details. Only students with high creativity scores managed this. The researcher also noted that creative individuals were more inclined to read, while those with lower creativity found reading difficult, struggled to visualize situations described in texts, had difficulty making inferences unless the author explicitly stated their ideas, and were more likely to believe the accuracy of information presented without critical observation.

Wang's (2012) study is another example investigating the link between language input and creativity. Wang aimed to explore whether there was any connection between students' attitudes towards reading and writing skills and their creativity, whether different reading and writing habits impacted their creativity, and how creativity scores of students who devoted more time to reading and writing compared to those who devoted less time. The study, involving 196 university students aged 18 to 21, used the Abbreviated Torrance Test for Adults (ATTA) to measure creativity. A questionnaire was used to gather information on students' attitudes towards reading and writing, the time they spent on these activities, and their background information. The findings revealed a positive correlation between students' reading habits and high creativity scores, particularly among those who discussed the books they read. Conversely, students without reading habits had lower creativity scores and did not favour reading only when necessary. Students who read for fun, read magazines or journals, and followed news online also exhibited high creativity scores. In terms of the time spent on reading, significant differences were observed as students did not all belong to the same department. Students from English and Chinese departments devoted more hours to reading and writing. Consequently, students from other departments such as information science and mathematics were found to spend less time on reading and writing compared to their peers from the English and Chinese departments. The ANOVA test indicated that students who devoted more time to reading and writing scored higher on creativity tests. Statistically, the most significant difference in creativity traits among the four departments was seen in elaboration, while originality and flexibility showed the least difference. Wang attributed this to the nature of elaboration: the higher the level of detail orientation in a student, the higher their elaboration score. Therefore, students in quantitative departments did not engage in rich verbal detail work while performing

specific numerical operations. Hence, reading and writing activities, which require high verbal intelligence and detail, explained the differences observed across departments. In summary, as reading and writing habits are developed and the time allocated to these activities increases, individuals' creativity also develops, thus recommending the cultivation of a positive attitude towards reading and writing.

A comprehensive study by Ritchie et al. (2013) involving 483 participants aged 15 to 23 holds significant importance in the context of the relationship between receptive skills and creativity. The researchers aimed to test whether reading difficulties could provide cognitive contributions to creativity, contrary to previous studies. While prior studies indicated that reading difficulties (such as dyslexia) were not linked with high intelligence and creativity, and was instead seen as distinct from general cognitive abilities both biologically and psychologically, there were quite a lot of studies that supported the idea of reading difficulties lead to higher creativity and intelligence. This particular study sought to disprove those hypotheses with a larger participant pool. To measure reading skills, the researchers used the CORE test, a 120-word test (with extra items for older participants). For creativity levels, the 12-item NEO-FFI test and the 48-item NEO-PI-R test were used. Sixty participants did not take part in the Queensland Core Skills Test (QCST), a routine test usually conducted at age 17, at the final year of their schools which includes a creative writing task. The findings indicated that the hypothesis suggesting individuals with reading difficulties would have higher creativity was incorrect. Instead, statistical data showed that individuals with higher reading scores also had higher creativity skills. These findings debunked the earlier hypothesis that individuals with reading difficulties had higher creativity and highlighted that creative skills correlated positively with both reading and creative writing abilities.

A similar study by Mourgues et al. (2014) also aimed to disprove the hypothesis that individuals with learning difficulties have higher creativity compared to normal individuals. Similar to Ritchie et al.'s (2013) study, the hypothesis tested was that students with lower verbal skills, particularly reading skills, would possess higher creativity scores. The study involved 259 first-year university students aged 17-38 from three different universities, grouped based on their majors: arts, social sciences, and applied sciences. Reading skills were measured using the Reading Comprehension task (RC), the Adult Reading History Questionnaire – Revised (ARHQ-R), and Compound Word Association tasks, while creativity scores were assessed using the Drawing Production

and Alternative Use Test (tests for phonological awareness and insight used in the study were not included as they were not relevant to the relationship between input and creativity). The findings supported Wang's (2012) study, revealing a positive correlation between students' reading skills and creativity levels, indicating that learning difficulties like dyslexia are not associated with higher creativity levels.

A recent study focusing on children instead of adults observed a similar setting with reading abilities and creativity. Bezerra et al. (2022) examined the relationship between creativity, intelligence, phonological awareness, and reading skills in 75 children aged 6-9 (first graders: 6-7 years, second graders: 7-8 years, third graders: 8-9 years). The Brazilian Figural Creativity Test (TCFI) was used to measure creativity, and the Reading Assessment of Words and Pseudowords Isolated (LPI) was used to assess reading skills. The study found that students' creativity scores increased slightly as they progressed from first to third grade. Researchers noted that since children at the elementary level go through different developmental stages, the slight differences were expected due to their similar ages and being in a similar literacy period. The study also mentioned that environmental factors such as teaching methods, material deficiencies, poor school conditions might have influenced the students' creativity. Although a correlation was found between reading skills and creativity as students aged, it was not significant in this particular study due to insufficient performance of the children in reading tasks. Consequently, even though a relationship has been found between the two variable similar to previous studies have shown, the relationship between reading skills and creativity could not be established as significant in this specific study.

One rare example of a study focusing on listening skills and creativity instead of reading skills is Karya et al.'s (2022) research. They aimed to observe whether the team games-based language learning (TG-BLL) model would enhance students' speaking and listening skills and whether students with high levels of creativity would also have advanced speaking and listening skills. In the study, 69 secondary school students were divided into two groups, with the experimental group learning through the TG-BLL model and the control group continuing their education with the existing, traditional, system. For this study, the second research question, "whether students with high creativity would also have high speaking and listening skills," is of particular importance. Students' creativity levels were categorized as "very high," "high," "medium," and "low." The results indicated that creative students performed better in listening skills than

less creative students. Although the specific creativity test used was not mentioned, the criteria for creativity scores were determined as the ability to create listening strategies, with students' task achievements used as a measure of creativity. Despite being one of the few studies directly comparing a receptive language skill with participants' creativity, it should not be considered valid or reliable.

2.1.3.2. Creativity and Language Production. As noted in many previous studies, creativity is most clearly observed in tasks that examine language production. Specifically, creative writing, which was studied as a concept even in the 1920s and 1930s, before the concept of creativity was considered a cognitive process in the 1950s, highlights this phenomenon (Leahy et al., 2014). However, as research began to explore the impact of creativity as a cognitive process on language skills, the definition of creative writing evolved over time. One of the major issues in defining creativity is the requirement for the product to be new and original, which also raised questions in creative writing. Consequently, researchers sought to determine which writings should be considered creative and for what reasons. Marshall (1974) specifically linked the creativity of written or spoken language to an individual's ability to convey personal experiences through their own symbolism. If individuals experience something directly or indirectly and then interpret and express it in their own way, such productions can be considered creative. Similarly, Dawson (2005) argued that creative writing need not be strictly literary and that any writing expressed in an unconventional manner could be considered creative. These definitions suggest that creative writing should be based on personal experiences and thus, involve unique symbolism (Wang, 2019).

In the literature, creative writing is often used either as a measure of creativity (Zhao, 2014; Wang, 2021), as a supplementary tool to foster creativity and motivation among students (Stillar, 2013; Arshavskaya, 2015; Yeh, 2017; Göçen, 2019), or to compare human creativity with artificial intelligence in contemporary studies (Roemmele, 2021; Ippolito et al., 2022; Bhimavarapu, 2023). However, studies that separately examine the cognitive, or social personality dimensions of creativity in the context of creative writing are scarce. This scarcity may be due to the inherent assumption that creative writing is creative by nature, thus negating the need to measure the creativity of the writers. However, albeit rare, there are studies in the literature that directly examine how creativity impacts writing skills.

For example, Nosratinia and Razavi (2016) designed a study to observe whether there was any relationship between students' creativity and their writing skills measured by complexity, accuracy, and fluency (CAF). The study involved 185 EFL students aged 20 to 32, using a Persian-translated version of O'neil et al.'s (1992) creativity scale. Students were asked to write two essays: one describing the characteristics of successful English teachers in Iran and another proposing solutions to unemployment. 50 minutes were allocated for each essay. To evaluate the essays based on CAF measurements, Larsen-Freeman's (2006) scale was used. The results indicated a strong positive correlation between creativity and each of the CAF measures. However, regression analysis showed that creativity had a similar impact on accuracy and fluency scores, while the effect on complexity scores was lower. As the authors noted, although creative individuals possess complex and diverse thoughts, it appears that accuracy and fluency in writing are more important to students than the complexity of the text.

In a similar study by Wati (2018), 80 students were divided into experimental and control groups to focus on writing skills that enhance creativity, testing whether these skills impacted students' creativity. Creativity was measured using the written form of the Torrance Test of Creative Thinking, while writing skills were assessed through a 90-minute essay. Statistical analyses revealed that the experimental group, which received instruction in creative writing, showed improved writing skills compared to the control group with similar creativity levels. Although a significant correlation consistent with previous studies was found, it is important to note that the criteria used to evaluate the essay in the research were not specified, raising questions about the study's validity.

In the study by Haerazi et al. (2020), the effectiveness of the genre-based instructional (GBI) model compared to the process-based instructional (PBI) model was examined by observing students' creativity skills. Sixty-four students were divided into experimental and control groups, with each group comprising students with high and low creativity levels. The students were asked to produce a narrative text, which was scored using a writing rubric to assess their writing skills. While the researchers mentioned that a verbal creativity test was administered to evaluate the students' creativity skills, they did not specify which creativity test was used. According to the results, the group taught using the GBI model demonstrated higher writing skills compared to the group using the PBI model. Additionally, students with higher creativity levels outperformed those with lower creativity levels in writing tasks. Another observation by the researchers was that

the instructional model had a more significant impact on creative students. Therefore, it was suggested that teaching methods should be chosen with an understanding of their substantial influence on students' creativity. Although Haerazi et al.'s (2020) study indicates that creativity positively affects writing skills and even influences the choice of instructional model, the lack of a specified valid and reliable creativity test raises questions about the study's reliability.

In a recent study by Bower et al. (2023), it was highlighted that evaluating students' creativity through their essays is an effective method, albeit time-consuming, thereby necessitating the testing of text-mining tools' assessment capabilities. Students aged 15-17 were given 40 minutes, and a total of 230 essays were collected. Each essay was assessed by four different teachers for creativity scores, quality, sophistication, semantic distance, originality, elaboration, idea density, the number of unique words, Shannon's entropy index, and total word count. In addition to the teachers' assessments, various tools were employed for comparison: MeanSim for semantic distance, Open Creativity Scoring for originality and elaboration, CPIDR5.1 for density, Planet Calc for unique words, Shannon's Entropy for entropy, and Linguistic Inquiry and Word Count tools for word count. The primary aim of the study was to explore the feasibility of automating these assessments by comparing teachers' evaluations with machine evaluations. The study's first research question, which is of primary importance here, focused on the relationship between the quality of students' essays and their individual creativity scores. The statistical findings indicated a direct positive correlation between both the quality and sophistication of the essays and the students' individual creativity scores. Machine evaluations aligned with teachers' assessments for each metric, except for the MeanSim results, which did not match the teachers' evaluations. These findings suggest progress in the capability of machine assessments to evaluate students' creativity. However, as the authors noted, further replication and validation of these assessments in different research contexts are necessary to ensure their reliability.

In D'Souza's (2021) systematic review, which included 39 empirical studies published between 2000 and 2020 examining how creativity is assessed in writing skills, it is noted that the research focus tends to be on creative writing rather than on the features of creativity termed "little-c"; creativity that is more aligned with everyday life and more widespread. However, creativity can be observed in all areas of life at any moment. As highlighted in definitions of creativity, the requirement for a product or idea to be new

and original does not imply that it is rarely seen or produced. D'Souza (2021) emphasizes that scientific research centred on creativity should examine both cultural and social dimensions of creativity. Therefore, it is insufficient to only address the concept of creative writing; it is necessary to consider creativity in all its aspects.

As previously mentioned, language production skills are divided into written and spoken language. However, since the primary focus of this study is on spoken production, it is addressed separately. A distinct section will discuss the differences between speaking and writing skills, the methods for measuring spoken production, and the empirical studies examining the relationship between spoken production and individual creativity.

2.2. Speaking

One of the most essential language skills is speaking. Long before humans discovered writing, they communicated with each other by making sounds. This ability, arising from the need to understand one another, has continued to develop intrinsically over time as the most fundamental communication skill. Florez (1999) states that speaking is an interactive process of constructing meaning that involves producing and receiving information both in form and meaning. Individuals, aware of the process of creating meaning, must determine what to say, when to say it, and why they say it. Considering these aspects before speaking is often an unconscious action for most people. Indeed, it is a fact that such processes are often unnoticed unless deliberately considered. As Thornbury (2005) notes, communication is a real-time activity that often lacks the time for detailed planning (p. IV). Moreover, communication frequently occurs in dialogue, requiring individuals to engage in multiple cognitive processes simultaneously in a fast-paced and dynamic manner (Johnson, 1996, p.55). This makes the process more complex than it might initially appear. The success of this process, which depends on various factors such as the communicative goals of the individuals, the social contexts in which they communicate, and their interlocutors, hinges on their communicative competence. Although communicative competence is a relatively recent concept, understanding its origin and significance for speaking is essential before delving into its definition and impact. The true importance of speaking among the four language skills has been recognized through this concept.

2.2.1. Speaking as a Language Skill

Throughout history, educators have developed educational models for language instruction based on the methods needed at that time. As noted by Brown and Lee (2015), until the 1980s, numerous researchers experimented with various methods in diverse contexts, cultures, and communities to find a single method suitable for each. In times when international cultural exchange was not as widespread as today, the primary goal of language learning was to follow the written literature in the target language(s) (Bailey, 2020). Therefore, the fundamental aim of language learning was centred on absorbing the grammar in texts and memorizing the words used in the text rather than learning them actively. This educational method, which did not prioritize speaking and listening skills, came to be known as the Grammar-Translation Method (GTM). However, as Richards and Rodgers (2014) pointed out, following the massive waves of migration in the 1950s after the two World Wars, the idea that education should have an international standard began to gain traction.

Due to these social and political changes, a transition phase known as The Reform Movement emerged to address the shortcomings of GTM. It questioned the effectiveness and durability of GTM in language teaching, thereby emphasizing the importance of listening and speaking skills. The initial foundations of communicative competence were laid with The Reform Movement. During this period of studying language acquisition processes, researchers examined how individuals acquire their native languages, leading to the conclusion that children develop a concept of language through listening before they start reading. In contrast to GTM, which focused on written texts, The Reform Movement initially prioritized listening by attempting to mimic children's habits with the sound systems of language. Unlike the emphasis on working with written texts, conversation practices where students would memorize or practice sentences and idioms gained importance (Larsen-Freeman & Anderson, 2011).

After The Reform Movement gained momentum, another method known as the Direct Method (DM) emerged as a primary focus for second language acquisition, mirroring the rapidly accelerating processes of native language acquisition. DM, which disallowed translation and encouraged teachers to foster direct, spontaneous dialogues within the classroom, became a system that encompassed all four language skills. It emphasized that the primary goal of language learning was communication, directing

even beginner-level students towards dialogue and providing correct pronunciation training (Bailey, 2020). Consequently, listening and speaking skills began to overshadow reading and writing skills, gaining greater prominence in educational practices.

During a similar period, in the mid-1900s, another instructional method that focused on speaking, known as the Audiolingual Method (ALM), was developed. ALM posited that habits were fundamental to individuals' language learning skills, emphasizing that communication occurred through specific spoken drills. Therefore, it argued that memorizing simple sentences or sentence structures in specific patterns was necessary (Thornbury & Slade, 2006). ALM frequently emphasized everyday conversations, yet instead of creating genuine conversation, students were expected to use scripted dialogues. Similarly, during a parallel timeframe, another theory known as Situational Language Teaching (SLT) was proposed. SLT, sharing similarities with ALM, teaches specific words and language structures that can be used in particular contexts. Both methods prioritize habit formation, hence placing significant emphasis on drill practice.

As seen from the era beyond the 1950s, when behavioural approaches were prevalent, educational methodological approaches diversified over time. In the 1960s, Total Physical Response (TPR) introduced by Asher (1966) emphasized listening and comprehension, aiming to measure whether students understood the prompts they listened to by physically responding to them. According to the TPR approach, it is asserted that if students can develop listening skills correctly, their speaking skills will naturally and spontaneously develop as well. TPR follows Krashen's (1982) Natural Approach, which underscores students' readiness, psychological states, and most importantly, teachers' ability to provide comprehensible input to students. Introduced by Curran in 1976, Community Language Learning posited that learning is more of a group effort than an individual endeavour, typically involving students recording their own voices for repeated listening and based on group discussions as an instructional method in history. Approaching the end of the 20th century, methods such as Suggestopedia, which emphasized the individual comfort of students, and Silent Way, which aimed to encourage students to speak without much direction from the teacher, were also experimented with, recognizing the importance of learning environments. However, the utilization of multiple methods revealed that rather than a singular method comprehensively encompassing all theoretical and practical aspects of language teaching,

it would be more accurate to emphasize the effective principles and procedures of various methods. Therefore, Kumaravadivelu (2006) proposed ten macro-strategies focusing on speaking and listening for students: a) maximizing learning opportunities, b) increasing negotiated interaction, c) minimizing perceptual mismatches, d) activating intuitive heuristics, e) enhancing language awareness, f) contextualizing linguistic input to students, g) integrating language skills, h) teaching and developing student autonomy, i) ensuring social relevance, and j) increasing cultural awareness.

After years of experimenting with different methods in various contexts, it has been recognized that language use skills are directly related to communication, hence to speaking and listening skills. When Hymes (1972) introduced the concept of communicative competence, he emphasized the need for individuals to know how to use language for various purposes in front of different listeners and in different social settings. He stated that this ability to use language is communicative competence (Duff, 2014, p.19). Communicative competence, which forms the basis of today's understanding, consists of four distinct components: a) linguistic competence (the ability of an individual to use the lexical and syntactic elements of language in speech), b) sociolinguistic competence (the ability of an individual not only to construct meaningful sentences but also to assess the appropriateness of what is said), c) strategic competence (the ability of an individual to plan communication, prevent or repair communication breakdowns), and finally d) discourse competence (the ability of an individual to produce coherent discourse using linguistic resources and sociolinguistic competence) added later (Goh & Burns, 2012, p.51) (Celce-Murcia, 2014).

Currently, the curriculum, materials, tests, and lesson plans in language education are focused on these four competencies. Therefore, based on the importance of these four competencies, the Communicative Language Teaching approach, which promotes real-life communication in the classroom environment, and also developed on not only linguistic accuracy but also linguistic fluency (Brown & Lee, 2015) has begun to be adopted by a wide range of educational communities. Therefore, due to the recognized significance of speaking skills among language competencies, there has been an increased focus on researching speaking skills in current studies compared to other competencies.

2.2.2. Two In/Different Modes of Language Production

Individuals' writing skills, like their speaking skills, are related to the productive functions of language. As Nan (2018) points out, each function of language indirectly or directly influences the others. While speaking skills rely more frequently on communicative competency, rhythm, intonation, and vocabulary knowledge as mentioned earlier, writing skills necessitate more analytical thinking and the use of more complex grammatical structures. Like other language skills, speaking skills also positively impact writing skills (or vice versa). Individuals' increased engagement in speaking allows them to interact more with linguistic material, thereby enabling them to use structures fluently in writing as well.

The differences between these two language skills, which contribute positively to each other, have been the main subject of numerous studies. Both cognitive (deliberateness and memory, emotion and attitude, and time and space) and linguistic (coherence, syntactic complexity) domains have revealed significant differences between them (Nicolau & Sukanto, 2018).

For example, in the study conducted by Lintunen and Makila (2014) with 18 Finnish students, they were asked to write essays of 150 to 250 words. The same students were also shown a series of 6 images and asked to describe the stories depicted in the visuals. Upon comparing these two sets of data, it was observed that although students exhibited more complexity in their written language, this difference widened further when a segmentation other than T-unit was preferred. In another study, Malvey (2017) instructed ninth-grade students to write persuasive compositions and later to orally share their ideas about working part-time while in high school. The results indicated that students used more complex syntax in their written texts. In contrast to these studies, Nicolau and Sukanto's (2018) study with second-grade Indonesian students revealed that while hedges, repetitions, and repairs were present in students' spoken language, their spoken narrative complexity scores were higher than their written narrative complexity scores. Researchers noted that repairs and repetitions observed in spoken language stemmed from students' attempts to speak rapidly.

Research comparing which is more difficult and complex between written production and spoken production has struggled to draw a consensus or reach a clear conclusion. Lintunen and Makila (2014) note that studies comparing these two skills view

both as cognitive tasks, suggesting that writing skills inherently require more cognitive processes and thus expend more cognitive resources. Furthermore, they point out that assessing writing skills entails higher expectations for accuracy due to the “thinking” time available during writing. In contrast, they emphasize that spoken production, by its nature, may include many inaccuracies in form, making it inappropriate to compare it directly in terms of accuracy.

Baron (2000) offers three distinct approaches to examine the difference between spoken production and written production:

- a. *The Opposition View*. Emphasizes that each production mode has distinct characteristics and cannot be evaluated from the same perspective. Written production relies on specific structures, is syntactically complex, and employs formal language. In contrast, spoken production lacks such rigid structures, features informal language, and syntactically exhibits much less complexity compared to written language.
- b. *The Crossover View*. Focuses on the diminishing distinction between different written and spoken varieties, where characteristics of one mode can increasingly be observed in the other.
- c. *The Continuum View*. Rather than separating production into spoken and written categories, proposes placing these modes along a single continuum. Academic and legal writings can be positioned closer to the literary end of this continuum, while informal correspondences lean towards the oral end. Therefore, the most significant differences are expected to be observed between varieties located at the opposite ends of this continuum when evaluated.

Consequently, studies comparing whether spoken production is more complex than written production or vice versa often select production varieties from the extremes of this continuum as criteria (Lintunen & Makila, 2014).

In this regard, as indicated by previous studies and numerous others (Leech et al., 1982; Brown & Yule, 1983; Pietila, 1999; Biber et al., 2002; Zhang, 2013), written production tends to incorporate a more complex and formal structure compared to spoken production. However, as highlighted in Halliday’s work (1979, 1989), spoken production can also exhibit complexity comparable to written language. Considering that the primary

purpose of writing is to be precise and concise, this distinction becomes clearer, as opposed to the simultaneous generation of ideas during speech production (Beaman, 1984).

In their study, Lintunen and Makila (2014) examine the syntactic complexity of written and spoken production to determine which is more complex. However, unlike previous studies, they use U-unit segmentation instead of AS-unit or T-unit for comparison. The researchers highlight that the fundamental issue lies here, noting that segmentation techniques suitable only for the complexity of written texts are used for comparing these two different modes, which do not align with the nature of spoken production. In their study comparing 18 essays and 18 spoken productions collected from Finnish upper secondary school students, it was observed that when both T-unit and AS-unit were used for comparison, written production significantly outpaced spoken production in terms of syntactic complexity. However, when segmented with U-unit, the difference was not as pronounced. Therefore, according to the findings of the research, although written production is generally more complex than spoken production, the gap starts to narrow when the segmentation unit is changed. Additionally, the researchers emphasize that individual differences heavily influence the results, noting that there are students whose spoken production appears more complex than their written production, contrary to general perceptions.

In a more recent study, Park (2022) aims to determine which mode—monologues or essays—exhibits greater syntactic complexity among Korean EFL students, analyzing 224 monologues and 139 essays. Building on previous research, this study attempts to address previous shortcomings by employing 14 different syntactic complexity indices. Of these indices, 11 showed higher values in writing, while 3 (coordinate phrases per clause, coordinate phrases per T-unit, and T-units per sentence) were higher in spoken production. It was noted that these 3 indices are related to coordination, whereas no significant differences were found among the remaining 11 indices. Therefore, although some differences were observed, it was found that these differences might be related to the selected segmentation unit, suggesting that these two production types may not be fundamentally distinct from each other.

2.2.3. *Measuring Speaking*

Effective communication in a foreign language facilitates individuals' integration into a new society/culture, removes barriers to academic success, and enables individuals to adapt more comfortably to society, thereby avoiding social isolation (Isaacs, 2016). Individuals must be able to express themselves successfully both orally and in writing in their second language. Although studies indicate that written production tends to exhibit, albeit marginally, more complex structures compared to spoken production, according to many researchers (Ur, 1996; Al-Sibai, 2004; Rabab'ah, 2005; Zhang, 2009; Burns, 2017; Pakula, 2019; Amoah & Yeboah, 2021), developing and advancing speaking skills tends to challenge students much more than improving their other skills.

Rababa'ah (2005) identifies several reasons why EFL students struggle with speaking English, categorizing them into issues stemming from students themselves, linguistic challenges, environmental factors, or teaching strategies. Misbah et al. (2017) suggest that the primary issue may stem from the influence of students' first language, inadequate vocabulary knowledge, or socioeconomic background. Ur (1996) asserts that student anxiety (fear of making mistakes or being criticized) significantly affects speaking skills, also noting that students may lack something to say or the motivation to speak. Similarly, Brown (2015) underscores that students' shyness and fears negatively impact their speaking abilities. In the context of Türkiye, Dil (2009) highlights anxiety and reluctance as the two most significant factors deterring students from speaking. Thus, anxiety and fear of criticism, as evidenced in numerous studies (Horwitz & Horwitz, 1986; Gebhard, 2000; Zhou et al., 2004), have a detrimental impact on students' speaking skills.

In order for students to fully demonstrate their spoken production skills, while the importance of classroom environment, surroundings, and educational materials is undisputed, it is crucial to accurately assess the performance students exhibit. Considering all these research findings, measuring whether students are correctly utilizing a skill in which they face such difficulties in production holds significant importance.

The initial approaches to measuring spoken production did not exclusively focus on spoken skills. During the second quarter of the 1900s, aural and oral skills were considered as a single skill. Despite efforts by figures like Lundeborg to emphasize

speaking as a separate skill, and pioneering initiatives such as the Certificate of Proficiency in English (CPE) in 1913 (Weir et al., 2013), the language learning paradigm of the time largely remained limited to grammatical correctness in translations, reflecting the Grammar-Translation Method (GTM). Until after World War II, assessment criteria predominantly evaluated students based on their writing and reading abilities. However, during the war, there arose a need to assess how well individuals could communicate in their second languages (Fulcher, 2003). The primary reason for this necessity, as identified by Kaulfer (1944), was the integration of military personnel into foreign societies for the gathering of essential military intelligence. Consequently, the earliest serious speaking assessment techniques emerged based on translation methods, albeit this time involving oral translations. These initial oral translations were particularly focused on information exchange and giving instructions, thus the criterion for successful language use was whether can do statements could be effectively utilized in specific communicative situations. Hence, this criterion appears to resemble closely the outcomes considered in tasks outlined in the Common European Framework of Reference (CEFR) today (Isaacs, 2016). Another subscale used by Kaulfer (1944), which continues to be prominent in oral assessment, is the concept of intelligibility. Evaluations of oral performance quality were conducted using a scale ranging from unintelligible to readily intelligible, reflecting a subjective assessment criterion that persists even in the absence of extensive research into oral assessment at the time (Isaacs, 2016). For instance, despite the commencement of the Test of English as a Foreign Language (TOEFL) in 1964, it consisted solely of reading, writing, and listening sections for the next 20 years.

During the predominant years of Structural Linguistics (from the 1920s to the 1970s), the development of separate assessment criteria specifically for speaking skills did not gain significant traction in the field of assessment. However, starting from the 1980s, with the shift towards communicative approaches in language teaching and testing, practices in foreign language education began to emphasize oral skills more frequently. Researchers and educators moved beyond the structural features of language to view language more as a tool for communication. Consequently, the objectives of developed programs were aligned accordingly, and language testers started evaluating language not as an isolated trait but within the dimensions of discourse, sociolinguistics, and contextual frameworks (Bachman, 2000). Thus, contemporary language teaching tasks are designed to be communication-oriented, preparing students with authentic

content that mirrors real-life scenarios, expecting them to achieve the communicative purpose of the task (Hatipoğlu, 2021). Additionally, De Jong (2023) emphasizes the necessity for tasks designed to assess foreign language skills to be valid, reliable, transparent, and authentic.

According to Fulcher (2014), tasks require students to produce a language sample that can be evaluated and scored. Alongside various task approaches (Johnson, 1985; Underhill, 1987; Weir, 1988; Bachman & Palmer, 1996), more recently, May (2010) has emphasized that tasks designed for assessing spoken production are increasingly specified in terms of context (the social setting where interaction occurs), interlocutor(s) (who communicates with whom and the social status relationship between communicators), and goals (the intentions of communicators, whether informative or interactional communication).

According to Pawlak (2016), types of tasks are categorized under five main headings, progressing from the most limited and non-spontaneous tasks to more spontaneous, demanding, and unrestricted tasks:

1. *Imitative Assessment Tasks*. As described by Brown and Abeywickrama (2004), these tasks involve students parroting back, being heavily monitored, and experiencing minimal spontaneous engagement. The primary and sole focus of imitative tasks is on specific phonetic components, where students listen and repeat without any meaningful exchange or contextual evaluation.
2. *Intensive Assessment Tasks*. Intensive tasks are limited and mechanical, requiring students to demonstrate specific and constrained lexical-semantic skills. Unlike imitative tasks, there is no interaction; examples include direct response tasks, oral readings, simultaneous sentence translations, and dialog completions.
3. *Responsive Assessment Tasks*. These tasks consist of simple question-answer pairs focusing on basic communication prompts such as requests, greetings, and related prompts. Initially, understanding and responding to questions are prioritized; at more advanced levels, students may be expected to ask information-seeking questions related to the prompt. Another example is paraphrasing complex prompts into simpler forms. However, research (Sun, 2009; McInnis, 2009; Milicevic & Tsedryk,

2011) indicates that students often struggle due to insufficient practice or inadequate lexical knowledge in such tasks.

4. *Interactive Assessment Tasks.* Interactive tasks aim to develop interpersonal; solidarity-building between interlocutors and learning how to determine the context of conversation and transactional interactions; teaching how information exchange occurs between interlocutors. Compared to other speaking tasks, interactive assessment tasks typically require longer and more complex responses. Examples include narrative tasks, comparison tasks, predictive tasks in oral interviews, role-plays, and discussions.
5. *Extensive Assessment Tasks.* These tasks focus on students planning, researching, and delivering in a more formal tone. Examples include oral presentations and storytelling, where students are expected to prepare in advance.

As mentioned, there are various methods for students to produce language samples. However, evaluating these samples based on criteria is a separate matter altogether. Ahmadi et al. (2020) note that assessment methods for evaluating oral skills generally fall into two categories: analytic scoring and holistic scoring.

Holistic scoring, as implied by its name, aims to provide a more general assessment using large-scale assessments, where evaluators consider predetermined criteria to evaluate performances. On the other hand, Taylor and Galaczi (2011) argue that the analytic approach is more beneficial for researchers because it evaluates each notable property of spoken production separately within its own subscale, providing more detailed information that can aid in planning for future learning and teaching purposes. Regardless of the method used, both approaches require scales and criteria to be established. When determining these scales and criteria, three main dimensions are often measured to evaluate students' foreign language performance: the complexity of the language used, the accuracy of the language used, and the fluency with which the language is used.

2.2.4. Complexity, Accuracy, and Fluency

In the 1970s, researchers sought new, reliable, and valid methods to observe how effectively students used language. The primary reason for this was, as mentioned earlier, the increased importance placed on oral proficiency and the shift towards observing spoken production rather than written production. Second language (L2) researchers aimed to develop an index that could quantitatively and verifiably measure students' performances. They decided to adapt grammatical complexity and accuracy metrics used in first language (L1) acquisition research to the L2 context (Housen et al., 2012). However, the use of these metrics pushed researchers in two different directions: one index used by L1 researchers known as mean length of utterance, and a second approach used within pedagogical frameworks that measured L2 performance in terms of accuracy and fluency dimensions (Michel, 2017).

Skehan (1998) innovatively integrated the three dimensions of L2 usage to create a new proficiency model: complexity, accuracy, and fluency (CAF). These concepts, whose original definitions are still considered valid today, have been used in numerous studies to observe the effectiveness (or ineffectiveness) of tasks as dependent variables. Through factor analyses, it has been recognized that CAF measurements are necessary for making general assessments about students' L2 performances or proficiency levels (Norris & Ortega, 2009). According to Housen et al. (2012), in recent years, alongside the cognitive approach in L2 research, CAF has become a central focus in its own right. As noted by Michel (2017), CAF variables are often used to manipulate independent variables. In this specific study, CAF measurements were employed to determine students' spoken proficiency levels. Given the multifaceted and multidimensional nature of CAF concepts inherent to the nature of language, it is crucial not only to define these concepts accurately but also to portray their relationships correctly.

2.2.4.1. Complexity. The construct of complexity within the CAF framework is perhaps the most debated aspect (Pallotti, 2015). According to Housen et al. (2012), complexity is divided into two main dimensions in the literature: linguistic complexity; the objective interpretation of complexity, focusing on the intrinsic formal or semantic-functional features of L2 elements, and cognitive complexity; the subjective difficulty of language features, or how learners perceive the difficulty of the items they process. Pallotti (2015) adds a third aspect, developmental complexity, which pertains to the

sequence in which linguistic structures emerge and are mastered. When measuring complexity in research, the linguistic dimension is typically employed (Michel, 2017). However, as noted by Bulte and Housen (2012), even the linguistic dimension is multifaceted, comprising lexical complexity and grammatical complexity as its subcomponents.

Jarvis (2013) categorizes the subdimensions of lexical complexity as follows: lexical diversity; the breadth of vocabulary, lexical sophistication; the depth of vocabulary, and lexical density; the amount of information in a text. Michel (2017) discusses grammatical complexity, noting that while syntactic complexity which entails sentence, clause, phrase structure is most commonly studied, morphological and phonological aspects should also be included in research. Moreover, he emphasizes the importance of selecting which components and how many for both lexical and grammatical complexity, as these choices significantly influence research findings.

In a similar vein, Norris and Ortega (2009) caution against selecting co-linear components and advocate for theoretical distinctions among different subcomponents. However, Inoue (2016) argues against this view, suggesting that task effects may be more critical than the choice of specific components. Lambert and Kormos (2014) similarly propose that these criteria may not be sufficient and advocate for different measures, such as distinguishing nominal subordination from subordination based on subject/object relative clauses.

Lastly, echoing Michel (2017), it is generally believed among researchers that the notion of more complexity leading to better language production is overly simplistic. Linguistic complexity varies even among different genres like small talk and argumentative tasks, indicating that higher complexity scores do not always correlate with higher performance levels. Therefore, in this study, following the aforementioned considerations, not only a single dimension was selected, but complementary dimensions of syntactic complexity and lexical complexity were examined, and their subcomponents were chosen according to the provided guidelines.

2.2.4.2. Accuracy. Accuracy, although perceived as the most observable construct within the CAF framework (Housen et al., 2012; Michel, 2017), inherently poses a significant problem. When broadly defined, accuracy encompasses any deviation from the norm in an L2 learner's written or spoken production (Hammerly, 1990; Pallotti,

2009). Such deviations are encoded as “errors” (Housen et al., 2012, p. 4). Despite the apparent clarity in identifying errors according to this definition and coding, the issue arises whether the linguistic norm should be a prescriptive grammatical standard or a target based on the native speaker (Michel, 2017, p. 54). According to Kuiken and Vedder (2014), opting for a prescriptive grammar as the norm may not be suitable for use within spoken production. Additionally, raters who perform these codings and ratings may not always reach a consensus on what constitutes accurate versus inaccurate (Michel, 2017), exemplified by varying normative standards within the German language across Austria, Germany, and Switzerland.

Moreover, Kuiken and Vedder (2008) propose a three-tiered assessment method for grading errors: “1” signifies a minor error of the first degree, “2” denotes a more serious error affecting comprehensibility, and “3” indicates an error rendering the sentence completely incomprehensible. More recently, Foster and Wigglesworth (2016) introduce the weighted-clause ratio, a system where students’ spoken or written productions are segmented into clauses, each rated on a scale where “1” denotes an error-free clause, “0.8” a clause with minor errors, “0.5” an error affecting meaning, and “0.1” a completely incomprehensible clause. Despite its initial appearance as a straightforward system aimed at quantitatively evaluating L2 performance, Foster and Wigglesworth (2016) caution that every rater focusing on accuracy will inevitably introduce personal judgments, making it nearly impossible to achieve entirely objective evaluations of L2 performance.

Michel (2017) notes that empirical studies frequently employ not only holistic scales and global measures but also specific measures. Holistic scales provide a general overview of errors, often without clearly distinguishing between complexity and accuracy, while global measures facilitate comparisons of errors across different languages, tasks, and participants but may not clearly differentiate minor differences at higher proficiency levels (Lambert & Kormos, 2014). Specific measures within accuracy reveal subtle differences, yet generalizing findings in relation to other research contexts remains challenging. In this study, the weighted-clause ratio method by Foster and Wigglesworth (2016) was chosen specifically to account for even minor errors in students’ spoken production.

2.2.4.3. Fluency. Fluency, a multidimensional concept like complexity and accuracy, focuses more frequently on spoken production compared to writing production

(Michel, 2017). However, there are established fluency measures also applicable to writing production. Fluency, often characterized as the smoothness or native-like quality of speech (Housen et al., 2012), is noted in research to sometimes be narrowly defined without encompassing all aspects of its definition (Lennon, 2000). Skehan's (2003, 2009) studies suggest that fluency can be categorized into three primary subdimensions: speed; frequency and rate of produced linguistic units, breakdown; number, length, and placement of pauses within produced linguistic units, and repair; false starts, self-corrections, repetitions, and misformulations (Housen et al., 2012; Michel, 2017).

According to Chambers (1997), fluency measures typically do not generate controversy among findings, especially when used within spoken production, thus generally considered reliable methods. Despite having defined subdimensions, de Jong et al. (2015) have shown in their studies that certain aspects of fluency are explained more by personality traits than by students' L2 skills. Therefore, instead of silence measures, they recommend using phonation time ratio which is defined as the percentage of speaking time relative to total speaking time. In light of all this information, while accuracy and complexity can be examined across all major levels of language—phonological, lexical, syntactic, and morphological—fluency is often viewed as a primarily phonological phenomenon due to its frequent and inherent nature.

In this study specifically, the relationship between students' creativity skills and their speaking proficiency was examined through the dimensions of fluency, speed (calculated by articulation rate), breakdown (pause durations calculated in two different dimensions in seconds), and repair (by measuring disfluency ratio).

2.3. L2 Speaking and Creativity

Due to the recent recognition of the importance of creativity in language learning, there is a lack of sufficient studies in the literature that directly examine how creativity “directly” affects language skills.

Otto's (1998) study is recognized as one of the pioneering works that examines the relationship between students' L2 proficiency and their creativity. The study involved 34 secondary school students aged between 14 and 16, and it measured participants' creativity through five different sub-tasks. In the consequences task, designed as a warm-

up exercise to prepare students for subsequent tasks, participants were asked to generate all possible outcomes of improbable scenarios (e.g., what would happen if all people suddenly became invisible, or if humans could suddenly speak the language of animals). The unusual uses task assessed students' originality by requiring them to list alternative uses for various objects (e.g., brick, nail, toothpick, wallpaper) beyond their intended purposes. In the common problems task, students were prompted to identify potential issues encountered in daily routines (e.g., going to school in the morning, making a sandwich). For ideational fluency, students were tasked with listing anything that came to mind within given categories (e.g., objects that are red or resemble red, such as "strawberry" and "raspberry"). Lastly, to measure associational fluency, students were instructed to complete three-word associations by linking two given words with a third word that was semantically related (e.g., "mirror" and "rain" associating with "water").

Students' language skills were assessed using both the Hungarian State Language Examination and their end-of-year grades. Otto (1998) noted that both measures of language proficiency consisted of tasks assessing communicative-based writing and speaking skills, which are considered the most reliable exam types for observing students' creativity and problem-solving abilities. The findings indicated a positive correlation between students' total creativity scores and their English language proficiency. This correlation remained consistent when individual task scores were examined separately. According to Otto's findings, students with higher creativity skills tended to perform better in language learning and usage compared to those with lower creativity skills. The researcher emphasized the importance of designing tasks that consider students' creativity levels, suggesting that tasks should accommodate less creative students as well. However, rather than adjusting tasks, it might be more reasonable to revise the curriculum to enhance students' creativity skills, as suggested by the researcher.

Although Otto (1998) indicated a strong relationship between language skills and creativity in his study, further research is needed to delve deeper into this relationship. Albert and Kormos (2004) investigate how students' creativity skills impact oral narrative tasks. Similar to Otto's (1998) study, their research involved 35 Hungarian secondary school students aged between 15 and 16. Participants completed two different oral narrative tasks two months apart, followed by the Hungarian creativity test. The creativity tests were divided into five sections. The first section involved a sentence completion exercise as a warm-up, which did not affect creativity scores. Following this warm-up,

students completed two oral narrative tasks. In the unusual uses task, students were asked to list unconventional uses for any given object (e.g., brick, nail, toothpick, wallpaper). The second task, defined as the distant associations task, required students to associate two given words with a third word that was semantically related. The last two tasks were adapted from the Torrance Tests of Creative Thinking (Torrance, 1966) and involved drawing tasks where students were instructed to draw freely without stopping from a circle, starting with any shape or image that came to mind. Additionally, picture narration tasks paired students to create stories based on a sequence of visual images. Each student had five minutes to prepare for these tasks, with one student narrating while the other assumed the role of the audience.

Creativity tasks were scored for fluency, flexibility, and originality. Total creativity scores were calculated by summing these scores, and subscores were tallied separately for total fluency, total flexibility, and total originality. For spoken production, linguistic measures included total words used, sentences per AS-unit, the ratio of correctly formed sentences to total sentences, lexical variety measured by the D-index, and the ratio of narrative clauses per AS-unit.

The findings indicated a moderate relationship between students' creativity skills and their performance on narrative tasks. Only 10-15% of linguistic measures correlated with creative fluency and originality. Overall, students who generated more solutions in creativity tests tended to speak more, thereby creating opportunities to use the target language. Students with higher originality scores spoke less but used a more complex language. Researchers emphasized that studies focusing on a single dimension of creativity may yield insufficient results, highlighting the need for research designed to encompass the multidimensional nature of creativity. They suggested that weak correlations could be attributed to the low participant numbers and noted that various social or individual differences could impact the results. Additionally, the researchers pointed out that granting students a long preparation time of five minutes might have limited the demonstration of creativity's full impact on the tasks.

In a longitudinal study, Smith (2013) discusses the expectation that students with higher creative skills should be more successful in classrooms where communicative teaching methods are predominantly used. His research aims to explore the connection between students' general English proficiency and their creativity, suggesting that creative students would likely perform better in instructional methods that involve open-

ended communicative tasks. Conducted with 58 first-year university students, the study divided participants into upper and lower groups based on placement tests conducted in the first half of the school term. Students' speech recordings from open-ended tasks conducted during lessons were transcribed by the researcher for analysis. To assess students' creativity skills, the researcher administered the Creative Behavior Inventory and a creative self-efficacy questionnaire.

As Smith (2013) notes, the study did not yield highly significant positive correlations; instead, moderate significant results indicated that speaking lessons focused on open-ended tasks were more beneficial for creative students. Therefore, the potential success of students with less creative skills in programs emphasizing such demanding tasks remains uncertain. Additionally, Smith (2013) highlights in his study that the creativity tests used may not fully reflect students' creativity skills, which should be considered when interpreting the absence of a meaningful correlation in the study.

In the study by Karimpour and Chopoglou (2014), the relationship between creativity and speaking skills was investigated using a sample of 62 students aged between 16 and 20. The study aimed to demonstrate that all students had similar language skills by initially administering the Comprehensive English Language Test, confirming that each student was at an intermediate level. Subsequently, the Arjmand Creativity Test was used to score students' creativity skills, and interviews were conducted within the classroom context to collect spoken data from students. During these interviews, students were expected to respond to common conversational prompts.

The findings indicate a strong positive correlation between students' speaking skills and their creativity. However, despite observing a positive correlation, it should be noted that the creativity test used in the study is a self-efficacy test, which may not fully reflect students' creativity across different dimensions and could potentially introduce bias. Additionally, questions remain regarding how the interviews were conducted and the expectations for students' use of speaking skills, which further necessitates a more detailed presentation of the research findings and a clear delineation of the methods employed in the study.

One of the significant studies on students' creativity and speaking skills is conducted by McDonough et al. (2015), aiming to determine whether students' creativity influences their performance in group problem-solving tasks and to what extent. The

study involved 55 first-year students at a Thai university. To measure students' creativity, the Torrance Tests of Creative Thinking A and B versions (picture construction task and picture completion task) were initially administered. Subsequently, a third task required students to complete lines or circles and create objects, followed by assigning titles to these objects. For the group problem-solving task, the "Shipwreck" scenario was employed, where four students in a group were tasked to decide which of the four survivors from a sinking ship should be sacrificed due to space limitations in a lifeboat.

Students' creativity scores were assessed based on subcategories such as fluency, originality, elaboration, abstractness of titles, and resistance to premature closure. Following previous studies, researchers opted to consolidate these sub-scores into a single creativity index. The conversations during the problem-solving task were recorded, transcribed, and compiled into a corpus of 11,321 words. Within this corpus, the most frequently used words were identified and categorized into six distinct groups: interactivity for questions, interactivity for subject/object pronouns, reasoning for subordinate clauses, reasoning for coordinated phrases, stance for verbs and modals, and predicting for conditionals.

The findings from the study, paralleling earlier research (e.g., Otto, 1998; Albert & Kormos, 2004), revealed a positive relationship between students' creativity and their ability to generate questions and coordinate ideas during the problem-solving task. However, no correlations were found with other linguistic features. According to the researchers, these results suggest that creativity allows students to vary their interactions during problem-solving tasks. Moreover, as highlighted by the researchers, the creativity measure used in the study primarily addresses the cognitive dimension of creativity, limiting the interpretation of the study's outcomes to a single dimension of creativity that necessitates a multidimensional approach.

Therefore, when evaluating students' creativity, researchers emphasize the importance of considering students' interests, attitudes, and past habits alongside cognitive assessments.

Wulandari (2016) aimed to investigate whether students with high creativity skills exhibit better speaking abilities compared to those with lower creativity skills. The study, involving 64 students, initially administered a 24-item creativity test developed by Munandar (2009). The findings from the research indicate that students with higher

creativity scores have higher average speaking scores than those with lower creativity scores. Therefore, based on these findings, the researcher claims that creativity has a positive impact on speaking skills. However, despite the apparent positive correlation within this study, there is no explanation provided regarding how the speaking tests were conducted and how the students' speaking results were evaluated. Additionally, due to the administered creativity test consisting of only 24 items, it is assumed to primarily assess creativity in terms of personality dimensions, suggesting a limited perspective on creativity as a multifaceted trait.

A study with a similar research design was conducted by Zuhriyah et al. (2018) involving 109 students. Initially, students were administered a written creativity test, followed by paired dialogues to collect spoken data. The creativity test used in the research consists of four main sections where students are required to generate a total of twenty sentences: five questions using the initial letters of given words, five questions using the middle letters of given words, five questions using the final letters of given words, and five additional questions using the last words of each of these fifteen questions. In the speaking test, after pairing students, tutors facilitated dialogues based on given conversation topics. The assessment of students' speaking skills primarily considered their comprehension abilities, vocabulary knowledge, grammar proficiency, fluency of speech, and pronunciation skills.

After completing the tests, students were divided into groups based on their high and low creativity levels, similar to the approach in Wulandari's (2016) study. Subsequently, Tukey tests were applied to analyse significant differences among the groups. The results indicate that, consistent with previous studies, students with higher creativity skills exhibited lower average speaking scores compared to those with lower creativity skills. Therefore, as noted by the researcher, students' creativity skills are directly related to and parallelly increase with their speaking skills. Additionally, the researcher highlighted that students with higher creativity are more motivated to learn and convey information, thus producing more diverse and meaningful ideas in speaking tasks, which they can articulate with broader vocabulary and correct grammar.

Although the results demonstrated a positive correlation, the creativity test administered to students, as acknowledged by the researcher, primarily assesses the divergent dimension of creativity, focusing on generating multiple alternative solutions to a situation. Therefore, while another study also found a significant relationship between

creativity and speaking skills, students' creativity was not fully examined in all its dimensions.

In Abdullah et al.'s (2020) study conducted with 30 English as a Foreign Language (EFL) students in Indonesia, the relationship between students' creativity and speaking skills was examined. To assess students' creativity, the Creative Personality Scale and Self-Rating of Personality tests were utilized, while students were required to deliver a monologue for their spoken data. The Creative Personality Scale consists of 30 items where students indicate their agreement on a scale scored from -12 to 18, assessing various aspects of creativity. The Self-Rating of Personality includes an 8-item section specifically focusing on Self-Rating of Creativity, where students evaluate themselves using Likert-type scales. For the spoken data, five items from the Creative Personality Scale were randomly selected, and students were asked to prepare a 5-minute monologue based on these selected items.

The findings of the study revealed a strong positive correlation between students' creativity scores and their speaking skills. However, similar to other studies, this research focused solely on one dimension of creativity (the personality trait dimension), without providing a comprehensive framework that includes its other dimensions.

As noted in several studies, there is a need for more detailed research on students' speaking skills. While the relationship between creativity and specific speaking tasks has been explored, albeit insufficiently, researchers particularly emphasize the necessity to investigate how creativity manifests under different task demands. In response to this gap, Suzuki et al. (2022) recently designed a study to observe which speaking tasks are more influenced by creativity. Building on previous research, Suzuki et al. (2022) focus on the stage of speech conceptualization, identified as the crucial point where creativity makes the most significant impact. They argue that the memory retrieval processes where individuals employ their creativity and the conceptualization processes used by students to prepare for speaking share similar cognitive operations, underscoring the importance of this stage for creativity.

Their study involved 60 students aged 18 to 23, who were tasked with two different speaking tasks involving distinct conceptualization process demands. The first task required students to generate their own ideas and arguments, thus involving a high demand on conceptualization processes, while the second task provided students with a

predefined speech content and required them to narrate events in a specific sequence. Creativity assessments for students included measures of DT, CT, and creative personality dimensions. Specifically, the Alternate Uses Test was used for DT, the Remote Associates Test for CT, and the Big Five Personality Test for assessing creative personality dimensions.

Regarding spoken data collection, one open-ended question and one picture narration task were employed. Students' spoken data were analysed for complexity (syntactic, lexical), accuracy (weighted-clause ratio), fluency (articulation rate, pauses, disfluency), and discourse features (total number of words, connectives). Additionally, placement exam results taken by students at the beginning of the semester were included to ensure students were at the same proficiency level.

The study examined a total of 15 linguistic measures and their relationship with students' placement exam results and three different dimensions of creativity. Findings indicated direct correlations between both divergent and CT with the use of connectives. Furthermore, DT showed a positive correlation with the total number of words in the argumentative task. In a broader perspective, creativity's cognitive dimensions were found to be associated with the discourse features of students' spoken data. Notably, a more significant relationship between creativity and spoken data was observed in the argumentative task compared to the picture narrative task, with researchers suggesting that the latter may constrain students' ability to speak freely as it guides them within a specific context. Thus, the correlation of DT observed in spoken data from picture narrative tasks suggests that students establish more unique relationships among events they perceive in visuals.

In conclusion, as highlighted by the researchers, creativity in both cognitive dimensions showed a relationship with conceptualization processes, regardless of low or high-demand contexts.

In the most recent study by Qin (2024), the direct impact of students' creativity and speaking skills on tasks is examined within narrative tasks focusing on two different task planning conditions. Conducted with 61 university students, the study randomly assigned students into two groups to ensure similarity in language proficiency levels using the Elicited Imitation Task. Subsequently, students completed a creativity test and its two sub-tests. A warm-up exercise preceded the narrative tasks, which were presented in two

versions requiring participants from both groups to narrate stories based on visual stimuli with and without planning time.

Creativity assessments in the study included the written version of the Alternative Uses Test, focusing on DT, and the circle drawing task from the Torrance Tests of Creative Thinking, which assessed creativity verbally and visually. Students' creativity was evaluated based on fluency, originality, flexibility, elaboration, verbal subscores, and figural subscores. Task performances were assessed in terms of complexity (average length of AS-units, number of sentences per AS-unit), accuracy (mean number of errors per 100 words), fluency (speech rate), lexical diversity (Vocd-D, measure of textual lexical diversity), quantity of talk (total number of words produced), narrative structure (temporal junctures linking independent clauses), and overall structure (ability to organize the story around a plotline rated on a scale of 1 to 10).

Analysis of the results revealed that students exhibited more fluent speech and increased productivity when given planning time, although no significant relationship was found between accuracy, complexity, and idea connection. An interesting finding was the negative predictivity of verbal originality, which the researcher attributes to limited planning time pushing students not only into the conceptualization process but also into formulating linguistic forms. This idea is linked to the observation that students construct more fluent and accurate sentences in planned tasks compared to unplanned tasks. In tasks without planning time, creative students were noted to narrate stories from visuals in a more detailed and imaginative manner, indicating that the retrieval of L2-specific vocabulary did not keep pace with idea generation and conceptualization processes.

Furthermore, proficiency tests conducted were found to correlate with task planning, indicating that more proficient students performed better in tasks with allocated planning time, thus suggesting the beneficial role of planning time for more proficient students.

In conclusion, the study highlights that students utilize their creativity most effectively in tasks without planning time, whereas tasks with planning time stimulate greater idea generation albeit less fluently. Supporting previous research, this study underscores that creativity serves as a facilitator for students within limited dimensions, emphasizing the need for further exploration in this area, as noted by the researcher.

SECTION III

METHODS

This section details the research design, participants, data collection instruments, data collection process, and the analysis of the collected data.

3.1. Research Design

The primary objective of the study is to observe how students' creativity skills influence speech conceptualization processes by collecting spoken data through one low-demanding speaking task and one high-demanding speaking task. Additionally, the study aims to investigate whether the preparation time for these tasks is directly or indirectly related to students' speaking proficiency or creativity skills by calculating the time taken for preparation. Due to the inability to exclusively collect data on the linguistic features of spoken language through qualitative methods, and with the objective of obtaining results from diverse numerical data, this study adopts a predominantly quantitative approach. Burns et al. (2015) define quantitative research as a "formal, objective, systematic process used to describe variables, test relationships between them, and examine cause and effect associations between variables" (Burns et al., 2015, p. 510). Therefore, quantitative research aims to test hypotheses using objective and impartial scientific methods (Davies & Fisher, 2018). Bloomfield and Fisher (2019) note that while there are many different research designs within quantitative research, fundamentally, it includes four different study designs. In this study, a correlational research design has been chosen.

Correlational research aims to discover whether there is a relationship between two or more variables and, if so, to explore the nature of this relationship. In correlational

studies, the variables are not manipulated, and the research typically focuses on examining the relationships between variables or testing theoretical models of these relationships rather than investigating cause-and-effect relationships (Shields & Smyth, 2016). Consequently, employing a research design such as qualitative design would be inappropriate for this study. This discrepancy arises from the fact that qualitative research primarily aims to comprehend and interpret social interactions, while quantitative research is concerned with hypothesis testing, examining causality, and making predictions, which are the main objectives of this study (Apuke, 2017).

3.2. Participants

Given the study's focus on English learners capable of conversing in English during class hours, the researcher opted to engage learners from his supervisor's class who were readily available for participation. The participants of the study consist of first-year students enrolled in the Speaking I course in the English Language Teaching department at Muğla Sıtkı Koçman University, who are taking this course for the first time. The 72 participants, aged between 17 and 24 ($M = 19$, $SD = 1.35$), include 47 female participants (65.27%) and 25 male participants (34.73%). The results of their placement test in the English preparatory program indicated that all participants could be classified as possessing B1 level proficiency, in accordance with the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2020). A background questionnaire further revealed that the majority commenced their English language education at the primary school level, as mandated in Türkiye, and had uniformly received primary and secondary education in the country.

In view of the temporal constraints inherent in the current study, rendering comprehensive coverage of the target population unattainable, the researcher decided to employ convenience sampling, a non-probability sampling method. Convenience sampling involves collecting data from participants who are readily accessible to the researcher (Rahi, 2017). Creswell (2012) posits that while these selected individuals may not represent the entire target population, they nevertheless yield valuable data for hypothesis exploration.

3.3. Data Collection Tools

The study incorporated three distinct creativity tests aimed at measuring creativity across three different dimensions (DT, CT, and openness to experience), along with the preparation of two distinct speaking tasks for spoken data. The tests designed to gauge creativity include the Remote Associates Test to measure CT, the Big Five Scale as the second to measure openness to experience, and the Alternate Uses Test as the third to measure DT. Regarding spoken data, the first task involved an argumentative question, while the second task utilized a picture narration task. In addition, a brief personal information form was also completed by the students.

3.3.1. *Personal Information Form*

First and foremost, participants completed a comprehensive personal information form (see Appendix 2), encompassing details such as their gender, age, self-reported proficiency level, and an open-ended query about their personal history as English language learners. This is due to reveal any potential connections between the results of the numerical data qualitatively.

3.3.2. *Remote Associates Test*

The Remote Associates Test (RAT) fundamentally serves as a psychological assessment that allows the evaluation of individual differences within associative processing. Developed by Mednick (1962), the RAT is frequently employed in contemporary research exploring constructs related to creativity, problem-solving, and memory (Lee et al., 2014). Given that the primary objective of the study is to comprehensively examine creativity and, in accordance with Guilford's (1967) definition, CT as the ability to find the most suitable solution to a problem among different ideas, the RAT is utilized to measure CT. Consequently, rather than subjecting students' creativity to a single test, a thorough examination is conducted from various perspectives. Instead of the abbreviated 33-item version of the RAT, a more comprehensive approach is adopted, utilizing the fully Turkish-translated and validated 120-item version by Özen et al. (2015) (Appendices 3, 4 and 5). The administration of the test involves presenting

students with three words on the smart board, allowing 15 seconds for consideration. Within this time frame, students are expected to identify a single word that, when added either at the beginning or end, completes all three words on the board. Students are anticipated to accurately complete all three words within the given time constraint. For example, if the words “elbise,” “oyunu,” and “ruhu” are simultaneously displayed on the board, the student is expected to identify that adding “takım” at the beginning of each word completes them. Similarly, a total of 120 words must be correctly identified, and students’ CT scores will be measured based on the number of correctly identified words out of the 120.

3.3.3. Big Five Scale: Openness

According to Simonton (2012), individual differences stand out as a characteristic trait, and thus, the act of creative thinking is observed as a more pronounced feature in creative personalities. A personality trait predisposed to manifesting creativity is an individual’s “openness to experience.” The personality analysis test that highlights this trait prominently is Goldberg’s (1992) Big Five Scale, consisting of five primary character traits (Neuroticism, Extraversion, Openness to experience, Agreeableness, and Conscientiousness). Although the Big Five Scale exists in various lengths and formats, the decision was made to use the original scale in its unaltered form to create a more comprehensive dataset. Cemalcılar et al. (2017) fully translated and validated this 60-item personality test into Turkish (see Appendix 6 for the inventory). Each item in the test begins with the statement “I see myself as someone who...” and includes adjectives such as “Outgoing, social, compassionate, kind-hearted.” Students are required to choose the degree of agreement with these adjectives on a scale of “1, Strongly disagree,” “2, Disagree a little,” “3, Neither agree nor disagree,” “4, Agree a little,” and “5, Strongly agree.” Only the “openness to experience” component of the test is considered as a dimension of creativity, and no data from the other personality traits has been utilized in the study.

3.3.4. Alternate Uses Test

In order to conduct a more comprehensive examination of creativity, both CT and openness to experience were investigated through the RAT and the Big Five Scale. DT,

which is one of the most frequently examined aspects of creativity (Ananiadou & Claro, 2009), was also explored. DT is succinctly defined as the ability to generate multiple alternative solutions to a problem, serving as the sole cognitive tool enabling individuals to think in multiple directions simultaneously (Kuhn & Holling, 2009). To observe students' DT abilities, Guilford's (1967) Alternate Uses Test (AUT) was selected as the most commonly used method. Following the underlying logic of the test, students are presented with the name of any object (e.g., brick) and are asked to continuously list, without pause, alternative uses for the object beyond its primary function (e.g., can be placed behind a door, under a table, used as a stove, etc.) for a duration of one and a half minutes. The test is administered entirely in Turkish, orally, and as time progresses, students can produce more creative and original ideas rather than ordinary alternatives that come to everyone's mind. The student's AUT scores will be measured based on the number of different words they can generate, and subsequently, these scores will be coded.

3.3.5. Argumentative Speaking Task

To gauge students' spoken proficiency levels, two distinct speaking tasks were devised for the collection of spoken data. The first of these tasks was an argumentative speaking task compromised of 40 independent speaking questions derived from the official TOEFL guide book (ETS, 2012) (see Appendix 7 for sample prompts). There are important reasons for the task being argumentative. Firstly, it was imperative to have a task that allows students to provide responses from their own perspectives (e.g. "Some students would prefer to live with roommates. Others would prefer to live alone. Which option would you prefer and why?" "Many universities now offer academic courses over the Internet. However, some people still prefer learning in traditional classrooms. Which do you think is better?"), moving beyond tasks that could lead to issues such as anxiety or stuttering in oral communication. Students were observed to produce much more relaxed and successful language when they could express their ideas from their own viewpoints (Maghfiroh & Hikmat, 2023). Another reason is that the questions, being answerable from the student's own perspective and open-ended, without any time constraints for answering or preparation, were motivating factors that reduced anxiety. According to Lambert et al. (2017), the choice of an argumentative speaking task is crucial for the integrity of the study due to its provision of a high-demand task type on

the speech conceptualization process, as it enables students to generate and select their own ideas.

3.3.6. Picture Narration Task

Despite the assertion by Lambert et al. (2017) that picture narrative tasks may be considered as having low conceptualization demand, Bailey (2005) contends that the incorporation of pictures represents one of the most effective methods for evaluating oral proficiency. This is primarily due to the tangible content pictures provide for discussion, diverting learners' attention away from themselves and thereby alleviating anxiety. Consequently, in this study, multiple picture narration (PN) tasks were employed to elicit students' speaking performance. The main objective of this task involves students examining a series of 6 to 8 pictures printed on A4-sized sheets (see Appendix 8 for the sample task), wherein the images sequentially depict a short story. All PN drawings used were taken from Campbell's (2022) doctoral dissertation. Students are expected to create a narrative based on these visuals and articulate the story orally. To ensure authenticity and prevent repetition for each student, eight distinct short stories were utilized.

3.4. Data Collection Procedure

Before commencing the data collection process, requisite approval was secured from the ethics committee of Muğla Sıtkı Koçman University (see Appendix 1 for the approval form). Due to the researcher's inability to be present at the university on specific days owing to physical constraints, the data collection period was condensed to two days. During the initial day, students were administered the Remote Associates Test (RAT), the Big Five Personality Test, and a personal information form. Considering the extensive nature of the RAT with its 120 items, it was deemed appropriate to conduct it on the first day, while the other two forms, being relatively brief, were sequentially completed on the same day. On the subsequent day, students were individually summoned through appointment scheduling to participate in individual sessions planned at 15-minute intervals. Both the researcher and the supervisor individually administered the AUT test to students, presenting them with an object and subsequently requesting alternative uses in sequence. Students, each allotted a time span of one and a half minutes for each object,

enumerated various uses to the best of their ability during this period. Following the AUT test involving four different objects, students were instructed to choose a number between 1 and 40. The corresponding independent speaking question associated with the selected number was orally presented to the student, who then had the freedom to utilize as much time as desired until ready to respond. Once prepared, the student responded without any time constraints. Following the argumentative task, students were again prompted to select a number between 1 and 8. The chosen story was displayed to the student on an A4 sheet, with a reminder that the student could examine the visuals for as long as desired. When ready, the student crafted and narrated the story without any time limitations. The entire process was meticulously recorded both audibly and visually to minimize data loss and subsequently transferred to digital media. No visual data were utilized in the research; solely the auditory data were transcribed for use within various analysis programs.

3.5. Data Analyses

Before commencing the data analysis process, the spoken data from students was transferred to the researcher's computer. Initially, all 144 spoken data samples were transcribed in their unpruned form (see Appendix 9 for sample transcription). This unpruned data was then segmented into AS-units following the guidelines provided by Foster et al. (2000), as a segmentation method that allows for a clearer analysis of spoken data (see Appendix 10 for sample transcription). Subsequently, based on the AS-units, a pruned version of the transcriptions was prepared (see Appendix 11 for sample transcription). The entire segmentation and transcription process was meticulously supervised and verified at each stage by an expert in the field.

For the analysis of the spoken data, the measurements established by Suzuki et al. (2022) were followed. For each CAF (Complexity, Accuracy, and Fluency) dimension, at least one representative large-grained measure was selected. In addition, to explore the relationship between students' creativity and speaking skills, several fine-grained lexical and discourse aspects of speech were also included as measures. The selected measures for each CAF dimension are as follows:

3.5.1. Complexity

In the study, the two major dimensions of complexity, namely lexical complexity and syntactic complexity, were examined separately.

3.5.1.1. Syntactic Complexity. Following Michel's (2017) guidelines, two large-grained measures were selected to assess the syntactic complexity of the students' spoken data:

1. *Subordination ratio.* Calculated as the average number of subordinate clauses per AS-unit.
2. *Clause length.* Calculated as the number of words per clause.

The researcher performed these syntactic complexity calculations manually without using any analysis software. The numerical results obtained were subsequently verified by another expert in the field.

3.5.1.2. Lexical Complexity. Following the study by Suzuki et al. (2022), the fine-grained measures selected based on the psycholinguistic properties of lexical items are as follows:

1. *Measure of textual lexical diversity (MTLD).* This measure, used to determine the variety of lexical items, is a variant of type-token ratio measures and is defined as “the mean length of sequential word strings in a text that maintains a given TTR value” (McCarthy & Jarvis, 2010, p. 384).
2. *Imageability.* This index, which indicates how easily individuals can visualize words in their minds, represents the averaged value of word imageability for all content words in the MRC psycholinguistics database (Wilson, 1998). Higher imageability scores correlate with a higher number of imageable words used in speech.
3. *Familiarity.* Defined as experience frequency, familiarity is calculated by averaging the word familiarity scores of all content words in the MRC psycholinguistics database. These scores, based on native speakers' subjective evaluations, are measured on a 7-point scale (1 = infrequently seen and used words, 7 = frequently seen and used words), multiplied by 100 to produce the final result.

4. *Hypernymy*. Defined as the specific meaning of a word in a sentence, hypernymy represents the average numbers of superordinate and subordinate words for all verbs and nouns in the WordNet database (Fellbaum, 1988). A lower hypernymy score indicates the use of more general and abstract words, rather than specific and concrete words, in speech.

To calculate the indices mentioned above, the researcher used the pruned transcriptions of the students' spoken data and computed them using the Coh-Metrix Web Analysis Tool developed by McNamara et al. (2014). Necessary permissions to use the Coh-Metrix tool, developed at Arizona State University, were obtained (see Appendix 12 for the approval mail).

3.5.2. Accuracy

As discussed in Section II, various methods are employed to measure the accuracy of students' spoken language. In this specific study, the measure referred to as the weighted-clause ratio by Foster and Wigglesworth (2016) was chosen. The pruned versions of the students' speeches were tabulated, and each sentence was individually scored. The scoring method varied according to the severity of the errors in the sentences, divided into four levels:

1. Error-free sentences = 1.0
2. Sentences with minor errors (e.g., using "go" instead of "goes," only missing an auxiliary verb, or incorrect use of singular/plural nouns) = 0.8
3. Sentences with serious but recoverable errors (e.g., errors in passive sentences, tense inconsistencies within the same sentence, or semantic errors due to incorrect word choice) = 0.5
4. Sentences with serious and unrecoverable errors, containing no meaningful content or lacking any verbs: 0.1 (Foster & Wigglesworth, 2016, p. 107)

After each sentence was scored, the total error score was divided by the total number of sentences to obtain the overall accuracy score for each student. The total accuracy score ranged from 0 to 1, with scores closer to 1 indicating higher accuracy (see Appendix 13 for sample ratings).

3.5.3. Fluency

Based on the work of Tavakoli and Skehan (2005), fluency is measured through three different subdimensions: speed, breakdown, and repair fluency. Four different fluency measurements were chosen for analyzing the discourse aspect of the students' spoken data:

1. *Articulation rate*. According to Sturm and Seery (2007), speech rate is directly related to effective communication; speaking too fast can overload the interlocutor with information, while speaking too slowly can interrupt the flow of information. Following the study by Suzuki et al. (2022), articulation rate was chosen to measure students' speaking speed. Articulation rate is calculated by removing all pauses from the students' recordings and then averaging the number of syllables per minute.
2. *Mid-clause pause duration*. Lambert et al. (2017) note that the location of pauses within speech can be used to identify differences in the underlying reasons for pauses, which is relevant for measuring breakdown fluency. Therefore, two pause-based measures were used for breakdown fluency, following Suzuki et al. (2022). The first measure, mid-clause pause duration, captures pauses within sentences.
3. *End-clause pause duration*. The second pause measure calculates pauses between sentences. Both pause measures were computed using the Praat program by Boersma and Weenink (2024) (see Appendix 14 for approval mail). Following previous studies (Suzuki et al., 2021), the pause duration was set at 250 ms. Automated detection of silence scripts was used to identify pauses, with necessary adjustments made by the researcher to determine the locations of pauses in the speech. Each pause was measured in seconds.
4. *Disfluency ratio*. Similar to the measurement of pause durations, the Praat program (Boersma & Weenink, 2024) was used to calculate the disfluency ratio. Disfluency features such as false starts, reformulations, corrections, and repetitions were identified and then divided by the total number of words spoken by the student. This ratio was multiplied by 100 to obtain the disfluency ratio.

3.5.4. Discourse

To comprehensively examine speaking skills, it is essential to consider discourse features in addition to the CAF (Complexity, Accuracy, Fluency) measurements. Following previous studies (Albert & Kormos, 2004; McDonough et al., 2015), Suzuki et al. (2022) highlighted the need to calculate the total number of words used by students to analyze the quantity of speech. Additionally, the use of connectives was emphasized due to its discovered correlations in these studies. Each measure was analyzed using the Coh-Metrix tool (McNamara et al., 2014):

1. *Total number of words.* The total count of all words produced by the student, excluding words involved in disfluency phenomena.
2. *Temporal connectives rate.* The rate of temporal connectives (e.g., before, after, later) used by the student, expressed per 1,000 words.
3. *Causal connectives rate.* The rate of causal connectives (e.g., so, since, because) used by the student, expressed per 1,000 words.
4. *Logical connectives rate.* The rate of logical connectives (e.g., and, but, or) used by the student, expressed per 1,000 words.

3.5.5. The Analysis of Creativity Measures

The first test used to measure creativity skills, the RAT, assesses students' CT abilities. The RAT consists of 120 items, and students' responses are measured by the number of correct answers out of these 120 items. The second creativity test, aimed at measuring students' social personality creativity skills, is the Turkish version (Cemalcılar et al., 2021) of the Big Five Inventory-2 (Soto & John, 2017), which is a shortened version of the Big Five Inventory (John, 1989; 1990). This personality inventory includes 60 items, of which 12 items related to openness to new experiences were used to calculate the students' openness scores.

The third and final creativity test, the AUT, measures students' DT abilities. The results of the AUT were initially analyzed by transcribing the responses from the audio recordings. The responses were then categorized to determine acceptable answers for the AUT. First, responses that included the primary uses of the objects were excluded. Then, responses serving the same function were grouped under the same category. After

categorizing the responses, the total number of responses given by each student was calculated to determine their AUT scores.

While DT tests typically categorize results under subdimensions such as fluency, originality, flexibility, and elaboration (Weiss & Wilhelm, 2022), and even though Suzuki et al. (2022) used originality as the most comprehensive subdimension for DT, this study specifically followed the guidelines of Dippo and Kudrowitz (2013) and opted to use only fluency. Thus, the total number of responses categorized under different purposes constituted the students' AUT scores.

After the transcription and segmentation processes were completed, all the collected data were entered into the Statistical Package for the Social Sciences 26 (SPSS) software for comprehensive statistical analysis.

SECTION IV

FINDINGS

In this section, the details and results of the statistical analyses conducted to answer the research questions are presented.

4.1. Predictor Variables

Table 1 presents the descriptive statistics for the predictor variables (placement test scores, DT scores, CT scores, openness scores) in the study.

Table 1. *Descriptive Statistics of Placement Test Scores and Creativity Variables*

	<i>M</i>	<i>SD</i>	<i>SE</i>
Placement Test Scores	79.97	7.88	.92
Divergent Thinking Scores	21.39	7.62	.89
Convergent Thinking Scores	64.89	12.75	1.50
Openness Scores	44.73	8.50	1.00
Preparation Time for AT	32.86	3.19	3.19
Preparation Time for PN	41.19	3.03	3.03

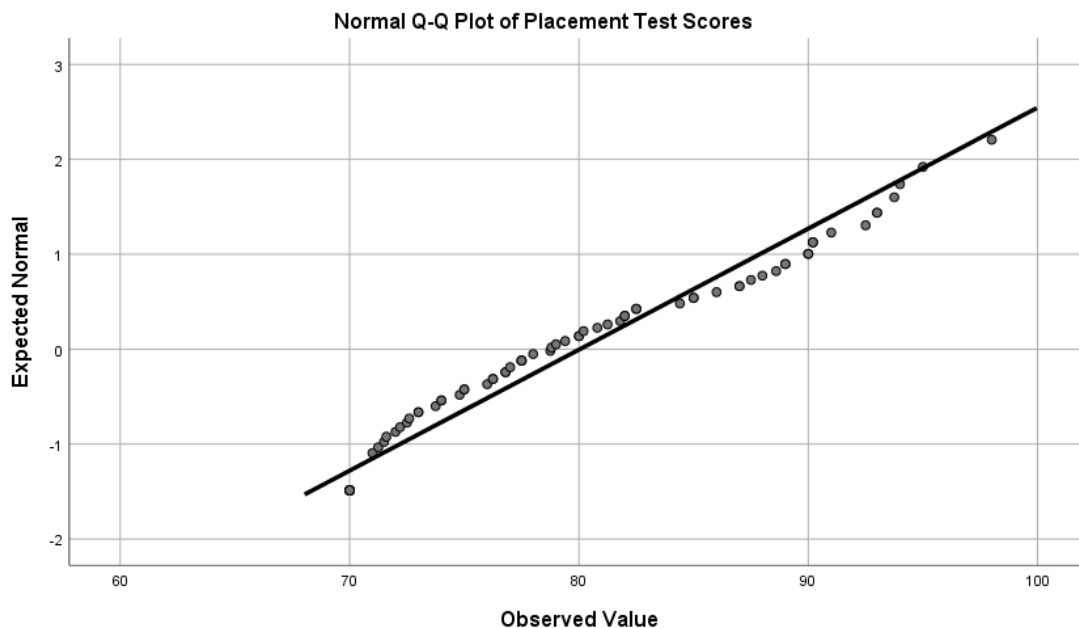
Note. AT = Argumentative Task; PN = Picture Narration

The majority of the students' placement test scores were measured using the Placement Test administered by the Prep School of Muğla Sıtkı Koçman University at the beginning of the 2023-2024 Fall semester. The results for a few students, however, were included in the calculations based on their scores from the YDS (Foreign Language Exam), prepared by the Council of Higher Education, which they took to be exempt from the Placement Test. Although the Placement Test scores students on all four language skills, those exempted through the YDS were assessed solely on their reading skills.

Therefore, it was necessary to examine whether the proficiency data of the students showed a normal distribution.

The normality of the data was assessed using Q-Q plots. As shown in Figure 1, the Q-Q plot for the placement test scores indicates that the data points closely follow the reference line, suggesting that the distribution of the data is approximately normal.

Figure 1. Q-Q Plot for Placement Test Scores



4.2. The Interrelationship Among Predictor Variables

Before presenting the results of the statistical analyses, the interrelationship between the predictor variables (i.e., DT score, CT score, openness to experience score) was examined to prevent potential multicollinearity in Table 2.

Table 2. Correlations of Creativity Variables and Placement Test Scores

	Divergent Thinking		Convergent Thinking		Openness Scores	
	r_s	p	r_s	p	r_s	p
Placement Test Scores	.016	.893	-.143	.232	.171	.152
Divergent Thinking	-	-	-.022	.852	.324**	.005
Convergent Thinking			-	-	.022	.857

Openness
Scores

- -

Note. ** $p < .01$

The interrelationship between the predictor variables (i.e., placement test scores, DT scores, convergent thinking scores, and openness scores) were examined due to preventing potential multicollinearity. As shown in Table 2, Pearson correlation (r) indicated that there were nonsignificant correlations between the placement test scores and the creativity variables. A negative correlation was found between DT and CT ($r_s = -.022$, $p = .852$), suggesting that DT and CT are two different cognitive aspects of creativity, yet, as it was stated previously, creative thinking processes comprise both. A significant positive correlation was found between DT and openness ($r_s = .324$, $p = .005$), indicating that DT and openness are both related to creativity, yet, former representing the cognitive perspective and latter representing the personality approach of creativity.

4.3. The Relationship Between Creativity and L2 Speaking Performance in AT

To answer the first research question, “Which predictor variables of creativity demonstrate significant associations with CAF measurements while L2 proficiency is controlled,” their creativity scores, and their placement test scores were examined. The descriptive statistics of students’ speaking performances in the AT are presented in Table 3.

Table 3. *Correlations of Creativity Variables and CAF Measurements in Argumentative Speaking Task*

	Placement Test		DT Scores		CT Scores		Openness	
	r_s	p	r_s	p	r_s	p	r_s	p
Subordinate Ratio	.070	.560	.200	.091	-.146	.222	.189	.112
Clause Length	.309**	.008	-.041	.729	-.228	.054	.077	.522
MTLD	.182	.126	-.072	.550	-.051	.671	.151	.204

Imageability	-.173	.146	-.055	.645	.042	.728	-.078	.513
Familiarity	-.245**	.038	-.082	.494	.148	.214	-.368**	.001
Hypernymy	.003	.980	.072	.548	-.170	.154	.095	.425
Accuracy	.324**	.005	.101	.398	-.054	.651	.076	.524
Articulation Rate	.452**	.000	-.076	.527	-.095	.429	.159	.182
Mid-clause pauses	-.364**	.002	-.055	.648	.039	.744	-.041	.732
End-clause pauses	-.165	.167	-.040	.739	.035	.773	-.171	.151
Disfluency	-.312**	.008	-.145	.226	.004	.975	-.201	.091
TNW	.381**	.001	.318**	.007	-.142	.234	.350**	.003
Temporal Connectives	.086	.473	.026	.826	-.041	.735	-.036	.765
Causal Connectives	-.090	.452	-.249*	.035	-.008	.946	-.087	.465
Logical Connectives	-.105	.379	-.211	.075	-.023	.848	-.158	.184

Note. * $p < .05$, ** $p < .01$; MTLTD = Measure of Textual Lexical Diversity; TNW = Total Number of Words

A correlation analysis was executed between CAF measurements, creativity scores and placement test scores in order to examine the function of creativity in argumentative speaking task. As shown in Table 3, Pearson correlation (r) indicated that the placement test scores and clause length ($r_s = .309$, $p = .008$), accuracy ($r_s = .324$, $p = .005$), articulation rate ($r_s = .452$, $p = .000$), and total number of words ($r_s = .381$, $p = .001$) were strongly positively correlated. While the placement test scores and familiarity ($r_s = -.245$, $p = .038$), mid-clause pause duration ($r_s = -.364$, $p = .002$), and disfluency ($r_s = -.312$, $p = .008$) were significantly negatively correlated. Amongst the creativity variables, DT scores and total number of words produced were significantly correlated

($r_s = .318, p = .007$), whereas DT scores and causal connectives were negatively correlated ($r_s = -.249, p = .035$). There were no significant correlations between CT scores and creativity variables. Openness was also significantly positively correlated with total number of words produced ($r_s = .350, p = .003$), and negatively correlated with familiarity ($r_s = -.368, p = .001$).

Considering the data in Table 3, a series of hierarchical multiple regression analyses were conducted to observe the predictive role of students' creativity skills on their speaking performance, while controlling for L2 proficiency. Only the three AT speaking performance variables that showed a significant correlation with at least one of the predictor creativity variables with weak effect sized established on Plonsky and Oswald's (2014) guidelines ($r_s > |.25|$) were included in the analyses. The statistical power required for the regression models ($N = 72$; maximum of four predictors) to detect a medium effect size was found to be .828, indicating that these models can be regarded sufficient in L2 research (Larson-Hall, 2010).

Table 4, 5, and 6 contain the results of the multiple regression analysis (MRA) of creativity skills and CAF measurements in AT.

Table 4. *Results of Hierarchical MRA Between Familiarity, L2 Proficiency and Openness in AT*

Model	R	R^2	SE	Change Statistics				Sig. F Ch.
				R^2 Ch.	F Ch.	df1	df2	
1	.245 ^a	.060	7.72	.060	4.48	1	70	.038
2	.411 ^b	.169	7.31	.109	9.03	1	69	.004

Note. a = Predictors: (Constant), Placement Test; b = Predictors: (Constant), Placement Test, Openness; c = Dependent Variable: Familiarity; Ch. = Change

Table 5. *Results of ANOVA^a Between Familiarity, L2 Proficiency and Openness in AT*

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	267.26	1	267.26	4.48	.038 ^b
Residual	4172.50	70	59.60		
Total	4439.77	71			
2 Regression	750.22	2	375.11	7.01	.002 ^c
Residual	3689.55	69	53.47		

Total	4439.77	71
-------	---------	----

Note. a = Dependent Variable: Familiarity; b = Predictors: (Constant), Placement Test; c = Predictors: (Constant), Placement Test, Openness

Table 6. *Coefficients of the Hierarchical MRA Between Familiarity, L2 Proficiency and Openness in AT*

	<i>Model</i>	<i>Unstandardized</i>		<i>Standardized</i>	<i>t</i>	<i>Sig.</i>
		β	<i>SE</i>	β		
1	(Constant)	607.75	9.34		65.04	.000
	Placement Test	-.246	.116	-.245	-2.11	.038
2	(Constant)	616.88	9.35		65.92	.000
	Placement Test	-.186	.112	-.185	-1.66	.101
	Openness	-.312	.104	-.335	-3.00	.004

Note. a = Dependent Variable: Familiarity

In order to determine the extent to which students' creativity skills explain the subdimension of familiarity within the CAF measurements, a multiple regression analysis (MRA) was employed. In this analysis, students' openness scores were used as the predictor variable while controlling for L2 proficiency (placement test scores).

Before presenting the results of the MRA, a number of assumptions were tested and controls were ensured. Firstly, stem-and-leaf plots and boxplots indicated that each variable included in the regression exhibited a normal distribution and did not contain univariate outliers. Additionally, the examination of the normal probability plot of standardized residuals and the scatterplot of standardized residuals against standardized predicted values demonstrated that the assumptions of normality, linearity, and homoscedasticity of residuals were met. Moreover, the Mahalanobis distance did not exceed the critical χ^2 for $df = 3$ (at $\alpha = .001$) of 16.27 for any case in the data, indicating the absence of multivariate outliers. Finally, all predictors showed relatively high tolerances, suggesting that multicollinearity did not impede the interpretation of the MRA results.

On the first step of the hierarchical MRA, since L2 was controlled, only the placement test scores were added and accounted for a significant %6 of the variance in familiarity, $R^2 = .060$, $F(1, 70) = 4.48$, $p = .038$. On the second step, openness was added to the regression equation, thus accounted for an additional %10.9 of the variance in familiarity, $\Delta R^2 = .11$, $\Delta F(1, 69) = 9.03$, $p = .004$. In combination, two predictor variables

explained %16.9 of the variance in familiarity, $R^2 = .17$, $F(2, 69) = 7.01$, $p = .002$. By Cohen's (1988) conventions, a combined effect of this magnitude can be considered "medium" ($f^2 = .20$).

Table 7, 8, and 9 display the results of the multiple regression analysis between the total number of words produced and the placement test scores, DT scores and openness scores.

Table 7. Results of Hierarchical MRA Between Total Number of Words Produced, L2 Proficiency, DT, and Openness in AT

Model	R	R^2	SE	Change Statistics				
				R^2 Ch.	F Ch.	df1	df2	Sig. F Ch.
1	.381 ^a	.145	113.14	.145	11.88	1	70	.038
2	.491 ^b	.241	107.36	.096	8.74	1	69	.004
3	.529 ^c	.280	105.36	.039	3.63	1	68	.061

Note. a = Predictors: (Constant), Placement Test Scores; b = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores; c = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores, Openness Scores; d = Dependent Variable: Familiarity; Ch. = Change

Table 8. Results of ANOVA^a Between Total Number of Words Produced, L2 Proficiency, DT, and Openness in AT

		Sum of	df	Mean Square	F	Sig.
Model		Squares				
1	Regression	152115.48	1	152115.48	11.88	.001 ^b
	Residual	896106.51	70	12801.52		
	Total	1048222.00	71			
2	Regression	252866.82	2	126433.41	10.96	.000 ^c
	Residual	795355.18	69	11526.88		
	Total	1048222.00	71			
3	Regression	293259.03	3	97753.01	8.80	.000 ^d
	Residual	754962.96	68	11102.39		
	Total	1048222.00	71			

Note. a = Dependent Variable: Total Number of Words Produced; b = Predictors: (Constant), Placement Test Scores; c = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores; d = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores, Total Openness score; Ch. = Change

Table 9. *Coefficients of the Hierarchical MRA Between Total Number of Words Produced, L2 Proficiency, DT, and Openness in AT*

	<i>Model</i>	<i>Unstandardized</i>		<i>Standardized</i>	<i>t</i>	<i>Sig.</i>
		β	<i>SE</i>	β		
1	(Constant)	-305.45	136.93		-2.23	.029
	Placement Test	5.87	1.70	.381	3.44	.001
2	(Constant)	-403.19	134.07		-3.00	.004
	Placement Test	5.77	1.61	.375	3.57	.001
	Openness	4.94	1.67	.310	2.95	.004
3	(Constant)	-470.20	136.19		-3.45	.001
	Placement Test	5.21	1.61	.338	3.23	.002
	Divergent Thinking	3.86	1.73	.242	2.22	.029
	Openness	3.01	1.58	.211	1.90	.061

Note. a = Dependent Variable: Total Number of Words Produced

To determine the extent to which students' creativity skills explain the total number of words produced within the CAF measurements, an MRA was employed. In this analysis, students' DT scores and openness scores were used as predictor variables while controlling for L2 proficiency.

As carried out earlier, before presenting the results of the MRA, a number of assumptions were tested and controls were ensured. As previously done, stem-and-leaf plots and boxplots indicated that each variable included in the regression exhibited a normal distribution and did not contain univariate outliers. Additionally, the examination of the normal probability plot of standardized residuals and the scatterplot of standardized residuals against standardized predicted values also demonstrated that the assumptions of normality, linearity, and homoscedasticity of residuals were met. Similarly, the Mahalanobis distance did not exceed the critical χ^2 for $df = 3$ (at $\alpha = .001$) of 16.27 for any case in the data, indicating the absence of multivariate outliers. Finally, all predictors also showed relatively high tolerances following the precedent set.

On the first step of the second hierarchical MRA, since L2 was controlled, only the placement test scores were added and accounted for %14 of the variance in TNW, $R^2 = .145$, $F(1, 70) = 11.88$, $p < .001$. On the second step, DT scores were added to the regression equation, thus accounted for an additional %9.6 of the variance in TNW, $\Delta R^2 = 0.10$, $\Delta F(1, 69) = 8.74$, $p = .004$. On the third step, openness scores were added to the regression equation, thus accounted for an additional %3.9 of the variance in TNW, ΔR^2

= 0.40, $\Delta F(1, 68) = 3.63$, $p = .061$. In combination, three predictor variables explained a significant 28% of the variance in TNW, $R^2 = .280$, $F(3, 68) = 8.80$, $p < .001$. By Cohen's (1988) conventions, a combined effect of this magnitude can be considered "large" ($f^2 = .38$).

Table 10, 11, and 12 display the results of the multiple regression analysis between the causal connectives rate and the placement test scores, and DT scores.

Table 10. Results of Hierarchical MRA Between Causal Connectives, L2 Proficiency and DT in AT

Model	R	R ²	SE	Change Statistics				Sig. F Ch.
				R ² Ch.	F Ch.	df1	df2	
1	.090 ^a	.008	27.52	.008	.572	1	70	.452
2	.264 ^b	.069	26.85	.061	4.54	1	69	.037

Note. a = Predictors: (Constant), Placement Test; b = Predictors: (Constant), Placement Test, Divergent Thinking Scores; c = Dependent Variable: Causal Connectives Rate; Ch. = Change

Table 11. Results of ANOVA^a Between Causal Connectives, L2 Proficiency and DT in AT

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	433.25	1	433.25	.572	.452 ^b
Residual	53045.70	70	757.79		
Total	53478.96	71			
2 Regression	3713.43	2	1856.71	2.57	.084 ^c
Residual	49765.52	69	721.24		
Total	53478.96	71			

Note. a = Dependent Variable: Causal Connectives Rate; b = Predictors: (Constant), Placement Test; c = Predictors: (Constant), Placement Test, Divergent Thinking Scores

Table 12. Coefficients of the Hierarchical MRA Between Causal Connectives, L2 Proficiency and DT in AT

Model	Unstandardized		Standardized		t	Sig.
	β	SE	β			
1 (Constant)	74.23	33.31			2.22	.029

	Placement Test	-.313	.415	-.090	-.756	.452
2	(Constant)	91.86	33.53		2.73	.008
	Placement Test	-.296	.405	-.085	-.731	.467
	Divergent Thinking	-.891	.418	-.248	-2.13	.037

Note. a = Dependent Variable: Causal Connectives Rate

In the third analysis, to determine the extent to which students' creativity skills explain the causal connectives rate within the CAF measurements, an MRA was employed. In the analysis, students' DT scores were used as a predictor variable while controlling for L2 proficiency.

Again, a number of assumptions were tested and controls were ensured and each variable included in the regression exhibited a normal distribution. Similarly, the Mahalanobis distance did not exceed the critical χ^2 for $df = 3$ (at $\alpha = .001$) of 16.27 for any case in the data. And all predictors showed relatively high tolerances.

On the first step of the third hierarchical MRA, since L2 was controlled, only the placement test scores were added but accounted for %0 of the variance in causal connectives rate, $R^2 = .008$, $F(1, 70) = .57$, $p = .452$. On the second step, DT scores were added to the regression equation, thus accounted for an additional non-significant %6.1 of the variance in causal connectives rate, $\Delta R^2 = 0.61$, $\Delta F(1, 69) = 4.54$, $p = .037$. In combination, two predictor variables explained a non-significant %6,9 of the variance in causal connectives, $R^2 = .070$, $F(2, 69) = 2.57$, $p = .084$. By Cohen's (1988) conventions, a combined effect of this magnitude can be considered "small" ($f^2 = .07$).

Under the next heading, the correlations related to students' creativity skills and their speaking performance in the PN task are shared, followed by the results of the regression analyses.

4.4. The Relationship Between Creativity and L2 Speaking Performance in PN

In order to answer the second part of the first research question, "In what ways does the task preparation time influence the quality of speech production and creativity," students' creativity scores, and their placement test scores were examined. The descriptive statistics of the students' speaking performances in the PN task are presented in Table 13.

Table 13. *Correlations of Creativity Variables and CAF Measurements in PN*

	Placement Test		DT Scores		CT Scores		Openness	
	r_s	p	r_s	p	r_s	p	r_s	p
Subordinate Ratio	.266*	.024	-.071	.556	-.212	.074	.051	.672
Clause Length	.435**	.000	.128	.282	-.215	.070	.232*	.050
MTLD	.416**	.000	.053	.657	-.002	.983	.052	.667
Imageability	-.168	.159	-.005	.970	-.123	.302	-.012	.922
Familiarity	-.066	.581	-.050	.677	-.072	.548	-.147	.218
Hypernymy	.116	.333	.094	.434	-.152	.202	.182	.126
Accuracy	.366**	.002	.025	.835	-.083	.489	.169	.157
Articulation Rate	.456**	.000	-.070	.558	-.194	.102	.091	.445
Mid-clause pauses	-.340**	.003	-.177	.137	.068	.571	-.043	.723
End-clause pauses	-.156	.189	-.283*	.016	.170	.152	-.265*	.024
Disfluency	-.068	.568	-.047	.693	-.027	.825	-.262*	.026
TNW	.279*	.017	.152	.203	-.133	.267	.217	.068
Temporal Connectives	-.051	.668	-.229	.053	-.038	.752	-.041	.732
Causal Connectives	.134	.261	-.121	.309	.145	.224	.047	.696
Logical Connectives	.064	.592	-.129	.280	.023	.850	.006	.957

Note. * $p < .05$, ** $p < .01$; MTLD = Measure of Textual Lexical Diversity; TNW = Total Number of Words

Another correlation analysis was executed between CAF measurements, creativity scores and placement test scores in order to examine the function of creativity in PN task. As shown in Table 13, Pearson correlation (r) indicated that the placement test scores and subordination ratio ($r_s = .266, p = .024$), clause length ($r_s = .435, p = .000$), MTLT ($r_s = .416, p = .000$), accuracy ($r_s = .366, p = .002$), articulation rate ($r_s = .456, p = .000$), and total number of words produced ($r_s = .279, p = .017$) were positively correlated. While the placement test scores and mid-clause pause duration ($r_s = -.340, p = .003$) was significantly negatively correlated. Amongst the creativity variables, DT scores and end-clause pause duration was negatively correlated ($r_s = -.283, p = .016$). There were no significant correlations between CT scores and creativity variables. Openness was found to be positively correlated with clause length ($r_s = .232, p = .050$), and negatively correlated with end-clause pause duration ($r_s = -.265, p = .024$), and disfluency phenomena ($r_s = -.262, p = .026$).

Considering the data in Table 13, a series of hierarchical multiple regression analyses were conducted to observe the predictive role of students' creativity skills on their speaking performance, while controlling for L2 proficiency. Only the three PN task speaking performance variables that showed a significant correlation with at least one of the predictor creativity variables with weak effect sizes established on Plonsky and Oswald's (2014) guidelines ($r_s > |.25|$) were included in the analyses. The statistical power required for the regression models ($N = 72$; maximum of four predictors) to detect a medium effect size was found to be .828, indicating that these models can be regarded sufficient in L2 research (Larson-Hall, 2010). The following tables contain the results of the multiple regression analysis of creativity skills and CAF measurements in PN task.

Table 14, 15, and 16 contain the results of the multiple regression analysis between clause length and the placement test scores, and openness scores.

Table 14. *Results of Hierarchical MRA Between Clause Length, L2 Proficiency and Openness in PN*

Model	R	R ²	SE	Change Statistics				
				R ² Ch.	F Ch.	df1	df2	Sig. F Ch.
1	.435 ^a	.189	3.50	.189	16.36	1	70	.000
2	.463 ^b	.214	3.48	.025	2.15	1	69	.146

Note. a = Predictors: (Constant), Placement Test; b = Predictors: (Constant), Placement Test, Openness; c = Dependent Variable: Clause Length; Ch. = Change

Table 15. *Results of ANOVA^a Between Clause Length, L2 Proficiency and Openness in PT*

	<i>Model</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
1	Regression	201.53	1	201.53	16.36	.000 ^b
	Residual	862.19	70	12.31		
	Total	1063.72	71			
2	Regression	227.69	2	113.84	9.39	.000 ^c
	Residual	836.03	69	12.11		
	Total	1063.72	71			

Note. a = Dependent Variable: Clause Length; b = Predictors: (Constant), Placement Test; c = Predictors: (Constant), Placement Test, Openness Scores

Table 16. *Coefficients of the Hierarchical MRA Between Clause Length, L2 Proficiency and Openness in PT*

	<i>Model</i>	<i>Unstandardized</i>		<i>Standardized</i>	<i>t</i>	<i>Sig.</i>
		β	<i>SE</i>	β		
1	(Constant)	-5.42	4.24		-1.27	.206
	Placement Test	.214	.053	.435	4.04	.000
2	(Constant)	-7.54	4.45		-1.69	.095
	Placement Test	.200	.053	.407	3.74	.000
	Openness	.073	.049	.159	1.46	.146

Note. a = Dependent Variable: Clause Length

In order to determine the extent to which students' creativity skills explain the subdimension of familiarity within the CAF measurements, a multiple regression analysis (MRA) was employed. In this analysis, students' openness scores were used as the predictor variable while controlling for L2 proficiency (placement test scores).

Before presenting the results of the MRA, a number of assumptions were tested and controls were ensured. Firstly, stem-and-leaf plots and boxplots indicated that each variable included in the regression exhibited a normal distribution and did not contain univariate outliers. Additionally, the examination of the normal probability plot of standardized residuals and the scatterplot of standardized residuals against standardized

predicted values demonstrated that the assumptions of normality, linearity, and homoscedasticity of residuals were met. Moreover, the Mahalanobis distance did not exceed the critical χ^2 for $df = 3$ (at $\alpha = .001$) of 16.27 for any case in the data, indicating the absence of multivariate outliers. Finally, all predictors showed relatively high tolerances, suggesting that multicollinearity did not impede the interpretation of the MRA results.

On the first step of the hierarchical MRA, since L2 was controlled, only the placement test scores were added and accounted for a significant %18 of the variance in clause length, $R^2 = .190$, $F(1, 70) = 16.36$, $p < .001$. On the second step, openness was added to the regression equation, thus accounted for an additional %2 of the variance in clause length, $\Delta R^2 = .25$, $\Delta F(1, 69) = 2.15$, $p = .146$. In combination, two predictor variables explained %21.4 of the variance in clause length, $R^2 = .21$, $F(2, 69) = 9.39$, $p < .001$. By Cohen's (1988) conventions, a combined effect of this magnitude can be considered "large" ($f^2 = .27$).

Table 17, 18, and 19 contain the results of the multiple regression analysis between end-clause pause duration and the placement test scores, DT scores and openness scores.

Table 17. Results of Hierarchical MRA Between End-clause Pauses, L2 Proficiency, DT, and Openness in PN

Model	R	R^2	SE	Change Statistics				
				R^2 Ch.	F Ch.	df1	df2	Sig. F Ch.
1	.156 ^a	.024	1.03	.024	1.75	1	70	.189
2	.320 ^b	.103	.99	.078	5.99	1	69	.017
3	.357 ^c	.128	.99	.025	1.96	1	68	.166

Note. a = Predictors: (Constant), Placement Test Scores; b = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores; c = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores, Openness Scores; d = Dependent Variable: End-clause Pause Length; Ch. = Change

Table 18. Results of ANOVA^a Between End-clause Pauses, L2 Proficiency, DT, and Openness in PN

Model	Sum of Squares		df	Mean Square	F	Sig.
	Model	Residual				
1	Regression	1.87	1	1.87	1.75	.189 ^b

	Residual	74.83	70	1.06		
	Total	76.71	71			
2	Regression	7.86	2	3.93	3.94	.024 ^c
	Residual	68.84	69	.99		
	Total	76.71	71			
3	Regression	9.79	3	3.26	3.31	.025 ^d
	Residual	66.91	68	.98		
	Total	76.71	71			

Note. a = Dependent Variable: End-clause Pause Length; b = Predictors: (Constant), Placement Test Scores; c = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores; d = Predictors: (Constant), Placement Test Scores, Divergent Thinking Scores, Openness Scores; Ch. = Change

Table 19. *Coefficients of the Hierarchical MRA Between End-clause Pauses, L2 Proficiency, DT, and Openness in PN*

	<i>Model</i>	<i>Unstandardized</i>		<i>Standardized</i>	<i>t</i>	<i>Sig.</i>
		β	<i>SE</i>	β		
1	(Constant)	3.16	1.25		2.53	.014
	Placement Test	-.021	.016	-.156	-1.32	.189
2	(Constant)	3.92	1.24		3.14	.002
	Placement Test	-.020	.015	-.151	-1.32	.191
	Openness	-.038	.016	-.279	-2.44	.017
3	(Constant)	4.38	1.28		3.42	.001
	Placement Test	-.016	.015	-.121	-1.05	.296
	Divergent Thinking	-.031	.016	-.225	-1.87	.065
	Openness	-.021	.015	-.171	-1.40	.166

Note. a = Dependent Variable: End-clause Pause Length

To determine the extent to which students' creativity skills explain the total number of words produced within the CAF measurements, an MRA was employed. In this analysis, students' DT scores and openness scores were used as predictor variables while controlling for L2 proficiency.

As carried out earlier, before presenting the results of the MRA, a number of assumptions were tested and controls were ensured. As previously done, stem-and-leaf plots and boxplots indicated that each variable included in the regression exhibited a normal distribution and did not contain univariate outliers. Additionally, the examination of the normal probability plot of standardized residuals and the scatterplot of standardized residuals against standardized predicted values also demonstrated that the assumptions of

normality, linearity, and homoscedasticity of residuals were met. Similarly, the Mahalanobis distance did not exceed the critical χ^2 for $df = 3$ (at $\alpha = .001$) of 16.27 for any case in the data, indicating the absence of multivariate outliers. Finally, all predictors also showed relatively high tolerances following the precedent set.

On the first step of the second hierarchical MRA, since L2 was controlled, only the placement test scores were added and accounted for %02.4 of the variance in end-clause pause length, $R^2 = .024$, $F(1, 70) = 1.75$, $p = .189$. On the second step, DT scores were added to the regression equation, thus accounting for an additional %10.3 of the variance in end-clause pause length, $\Delta R^2 = 0.08$, $\Delta F(1, 69) = 5.99$, $p = .017$. In the third step, openness scores were added to the regression equation, thus accounting for an additional %2.5 of the variance in end-clause pause length, $\Delta R^2 = 0.025$, $\Delta F(1, 68) = 1.96$, $p = .166$. In combination, three predictor variables explained a non-significant %12.8 of the variance in end-clause pause length, $R^2 = .128$, $F(3, 68) = 3.31$, $p = .025$. By Cohen's (1988) conventions, a combined effect of this magnitude can be considered "medium" ($f^2 = .14$).

Table 20, 21, and 22 display the results of the multiple regression analysis between the disfluency phenomena and the placement test scores, and openness scores.

Table 20. Results of Hierarchical MRA Between Disfluency, L2 Proficiency and Openness in PN

Model	R	R^2	SE	Change Statistics				
				R^2 Ch.	F Ch.	df1	df2	Sig. F Ch.
1	.068 ^a	.005	12.91	.005	.329	1	70	.568
2	.263 ^b	.069	12.58	.064	4.77	1	69	.032

Note. a = Predictors: (Constant), Placement Test; b = Predictors: (Constant), Placement Test, Openness; c = Dependent Variable: Disfluency; Ch. = Change

Table 21. Results of ANOVA^a Between Disfluency, L2 Proficiency and Openness in PN

Model	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	54.85	1	54.85	.329	.568 ^b
Residual	11679.53	70	166.85		
Total	11734.38	71			
2 Regression	810.44	2	405.22	2.56	.085 ^c

Residual	10923.93	69	158.31
Total	11734.38	71	

Note. a = Dependent Variable: Clause Length; b = Predictors: (Constant), Placement Test; c = Predictors: (Constant), Placement Test, Openness Scores

Table 22. *Coefficients of the Hierarchical MRA Between Disfluency, L2 Proficiency and Openness in PN*

	<i>Model</i>	<i>Unstandardized</i>		<i>Standardized</i>	<i>t</i>	<i>Sig.</i>
		β	<i>SE</i>	β		
1	(Constant)	15.23	15.63		.975	.333
	Placement Test	-.112	.195	-.068	-.573	.568
2	(Constant)	26.65	16.10		1.65	.102
	Placement Test	-.036	.193	-.022	-.188	.851
	Openness	-.390	.179	-.258	-2.18	.032

Note. a = Dependent Variable: Disfluency

In the third analysis, to determine the extent to which students' creativity skills explain the causal connectives rate within the CAF measurements, an MRA was employed. In the analysis, students' DT scores were used as a predictor variable while controlling for L2 proficiency.

Again, a number of assumptions were tested and controls were ensured and each variable included in the regression exhibited a normal distribution. Similarly, the Mahalanobis distance did not exceed the critical χ^2 for $df = 3$ (at $\alpha = .001$) of 16.27 for any case in the data. Hence, all predictors showed relatively high tolerances.

On the first step of the third hierarchical MRA, since L2 was controlled, only the placement test scores were added but accounted for %0 of the variance in causal connectives rate, $R^2 = .005$, $F(1, 70) = .57$, $p = .329$. On the second step, openness scores were added to the regression equation, thus accounting for an additional non-significant %6.9 of the variance in causal connectives rate, $\Delta R^2 = 0.70$, $\Delta F(1, 69) = 4.77$, $p = .032$. In combination, two predictor variables explained a non-significant %6.9 of the variance in causal connectives, $R^2 = .070$, $F(2, 69) = 2.56$, $p = .085$. By Cohen's (1988) conventions, a combined effect of this magnitude can be considered "small" ($f^2 = .07$).

4.5. The Impact of Preparation Time on L2 Speaking Performance and Creativity

In order to answer the second research question, “In what ways does the task preparation time influence the quality of speech production and creativity,” students’ preparation time prior to the tasks, their speaking performance, and their speaking performance were examined. The descriptive statistics of students’ preparation time in both tasks are presented in Table 23.

Table 23. *The Descriptive Statistics of Students’ Preparation Time*

	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>Min.</i>	<i>Max.</i>
PT for AT	32.86	27.12	3.19	4	140
PT for PN	41.19	25.77	3.03	7	158

Note. Preparation times are provided in seconds; PT = Preparation Time; AT = Argumentative Task; PN = Picture Narration

To observe whether there is any relationship between students’ preparation times for tasks and their creativity skills, a correlation analysis was first conducted between creativity skills and preparation times. The results of the conducted analysis are presented in Table 24.

Table 24. *Correlations of Creativity Variables and Preparation Time*

	AT Preparation Time		PN Preparation Time	
	<i>r_s</i>	<i>p</i>	<i>r_s</i>	<i>p</i>
Divergent Thinking	-.170	.154	-.067	.577
Convergent Thinking	.045	.709	-.006	.961
Openness	-.073	.545	-.042	.728

Note. AT = Argumentative Task; PN = Picture Narration

As can be seen in Table 24, no significant relationship was found between students’ creativity skills and their preparation times for the speaking tasks. Therefore, Table 25 presents a correlation analysis examining whether there is any relationship between students’ preparation times for the speaking tasks and their speaking performance in argumentative task.

Table 25. *Correlations of CAF Measurements and Preparation Time*

	Preparation Time in AT	Preparation Time in PN
--	------------------------	------------------------

	r_s	p	r_s	p
Subordinate Ratio	.031	.796	-.010	.934
Clause Length	-.107	.372	-.044	.716
MTLD	-.193	.105	.012	.921
Imageability	.052	.662	.119	.320
Familiarity	.075	.532	.198	.095
Hypernymy	-.089	.459	.034	.775
Accuracy	-.173	.146	.028	.818
Articulation Rate	-.333**	.004	.013	.916
Mid-clause pauses	.135	.260	-.060	.614
End-clause pauses	.058	.627	.096	.422
Disfluency	.174	.143	-.055	.645
TNW	-.113	.343	.047	.693
Temporal Connectives	.208	.079	.035	.770
Causal Connectives	.185	.121	-.259*	.028
Logical Connectives	-.002	.989	-.140	.240

Note. * $p < .05$, ** $p < .01$; MTLD = Measure of Textual Lexical Diversity; TNW = Total Number of Words

A correlation analysis was executed between CAF measurements and preparation times in both tasks in order to examine the function of preparation time in speaking performance. As shown in Table 25, Pearson correlation (r) indicated that the preparation time in AT and articulation rate has correlated negatively ($r_s = -.333$, $p = .004$). No other correlation was observed between students' preparation times for AT and the CAF measurements. Another correlation analysis was conducted to examine the relationship between students' preparation times for the second task, PN, and their performance on this task. According to the analysis results, a negative correlation was observed between causal connectives and students' preparation times for PN ($r_s = -.259$, $p = .028$). No other significant relationships were found in the results of the analysis.

SECTION V

DISCUSSION, CONCLUSION, and SUGGESTIONS

5.1. Discussion

In this section, the results of all the data collected for the research are explained under separate headings, in line with the research questions. The impact of students' creativity skills on their L2 speaking performance was observed through two different speaking tasks. Following this, the potential relationship between the preparation times for these tasks and students' creativity or speaking performance was examined. The similarities or differences between the findings and previous studies are discussed, and the results are explained in detail.

5.1.1. Discussion for the First Research Question

To address the first research question of the study, “Which predictor variables of creativity demonstrate significant associations with CAF measurements while L2 proficiency is controlled,” data on both students' creativity and L2 speaking performances were collected. To assess students' creativity, the AUT, RAT, and BFI tests were administered, while for evaluating speaking performance, an open-ended argumentative speaking task and a PN task were used. Analysis of the collected data revealed that the dimensions of DT, CT, and openness exhibited a normal distribution. Descriptive statistics of the data showed that the students' DT scores, measured by the AUT, averaged 21. In comparison, previous studies have reported average alternative uses of 10 in Dippo and Kudrowitz's (2013) AUT test with 293 participants, 13 in Lynch's (2015) study with 189 participants, 11 in Anbe's (2016) study with 524 participants, and 6 in deChantal and

Markovits' (2017) study with 96 participants. Despite the relatively small sample size in this study, students' DT scores were found to be high compared to these previous studies.

The correlation analysis between students' DT scores and the CAF measurements of their AT results revealed a positive correlation between the total number of words produced and DT, while a negative correlation was found with causal connectives. In line with the studies by Suzuki et al. (2022) and Qin (2024), an increase in a student's DT skills correlates with a higher likelihood of finding alternative solutions to a given problem. The findings indicate that more creative students generate more ideas, enabling them to express themselves more clearly by producing a larger number of words in response to a prompt. The results of the MRA demonstrate that when both L2 proficiency and DT scores are included in the analysis, they account for 24% of the variance. Therefore, both linguistic competence and DT skills positively impact the number of words used in a student's L2 speaking performance, consequently influencing the number of ideas generated.

A negative correlation was found between DT and the use of causal connectives, which are used to establish cause-and-effect relationships. In open-ended questions, particularly those designed to allow unrestricted idea generation, students are expected to be aware of the relationships between propositions in their speaking performance. However, despite Suzuki et al.'s (2022) findings indicating that cohesive devices and creativity skills affect speaking performance, similar results were not obtained in this study. The primary reason for this discrepancy may be attributed to the collection of spoken data within the context of the Speaking I course, where students might prioritize the quantity of talk over content due to concerns about grades. It can be suggested that in speaking performances with frequent use of connectives, students place greater emphasis on discourse-level organization, preferring clear and specific structures over establishing unique relationships between events or presenting more creative ideas during the conceptualization process. In conclusion, although the MRA results indicate that the use of causal connectives explains 6% of the variance without yielding significant results, this study did not establish a positive correlation between connectives and predictor variables.

Another aspect of students' cognitive creativity skills, CT, is assessed using a 120-item test. Students successfully answered only half of the test, averaging 64 correct responses. Consequently, while the DT dimension of students' cognitive creativity indicates high levels of creativity, the CT dimension represents an average level.

Although Özen et al. (2015) translated and validated the RAT test into Turkish, it was noted that the 120-item RAT test was excessively long and exhausting for participants, potentially affecting the results. Therefore, it was recognized that shorter versions of the test should also be translated and validated. Due to these reasons, no correlation was observed between CT and either of the tasks. Hence, within this study, only the DT dimension represents the cognitive aspect of creativity.

To represent the social personality dimension of students' creativity, openness, one of the sub-personality dimensions of the BFI, was selected. In the BFI test, which has a maximum score of 60, students averaged 44 points for openness. A correlation analysis for the interrelationship between DT and openness revealed a significant relationship ($r_s = .324, p = .005$), indicating that although these dimensions represent different aspects of creativity, both predictors clearly reflect creativity.

Considering the correlation results within the AT, openness was found to have a significant negative correlation with familiarity. Familiarity indicates how unique the words used in students' speaking performances are. The familiarity score ranges from 1 to 7, with lower scores indicating more unique words. Contrary to the findings of Suzuki et al. (2022), the correlation results show that as students' openness scores increase, the uniqueness of the words they use in their sentences also increases. Therefore, it can be inferred that the more open students are to new experiences, the more inclined they are to use new and unique words in their sentences.

Openness also shows a significant correlation with the total number of words used in students' speaking performances during the AT. According to the results of the MRA model, DT accounts for 24% of the variance in the total number of words produced, while openness contributes an additional 4%. Therefore, both openness and DT together significantly explain 28% of the total word count. This indicates that individuals who are more cognitively creative and those who are open to new experiences are more inclined to generate more ideas compared to less creative individuals.

The significant relationships between the CAF measurements of students' speaking performance in the AT and their creativity skills were examined. The next section will discuss the significant relationships between speaking performance in the PN task and creativity skills.

A negative correlation was observed between end-clause pause duration and the CAF measurements of speaking performance in the PN task with DT and openness. DT and openness together explain 12% of the variance, indicating that as students' creativity skills increase both cognitively and as a personality trait, the pause times between sentences decrease. One possible reason for this is the predefined content presented by the task itself. As students navigate through the images to create a narrative, they often examine unique relationships between events and characters. Thus, although Suzuki et al. (2022) attributed the relationship between DT and the PN task to the use of logical and causal connectives, the reduction in waiting time between sentences can similarly be attributed to DT skills. While openness showed a significant relationship, the MRA results indicate that openness contributes only a nonsignificant 2% to the regression model, suggesting that students' cohesion awareness is generally explained by DT in this model. Qin (2024) observed that as students' creativity levels increased, they were able to produce more ideas, but their speaking fluency decreased. However, in this study, another significant relationship from the PN task data was between openness and clause length. Contrary to Qin's (2024) findings, the speaking performance from the PN task in this study showed that students spoke fluently even while constructing longer sentences with well-formed syntactic structures in line with the ideas they wanted to express. Additionally, openness showed a significant negative correlation with disfluency phenomena, indicating that students with higher creativity skills are fluent and accurate when retrieving and articulating syntactic structures.

Suzuki et al. (2022) found a significant relationship between openness and subordination ratio in PN tasks. However, in this study, although students constructed longer sentences, no significant relation was found between subordination ratio and creativity skills. The use of subordination involves establishing a hierarchical relationship between sentences, with each sentence built around a single proposition (Kormos, 2006). Despite constructing longer sentences, students were unable to create different sentences based on a single proposition. The low CT scores of the students could also be a reason for this. Although it has been suggested that a shortened and validated version of the CT test is needed, considering that the test results accurately reflect the students' CT skills, it is plausible that the students were unable to use subordinate clauses. Given their CT skills, students should have been able to connect different prompts to a single proposition to arrive at a single outcome.

Suzuki et al. (2022) also reported a significant relationship between DT and the use of causal and logical connectives in PN tasks. However, this study, in line with Qin's (2024) results, did not find any significant data between connectives and creativity skills. Qin (2024) suggested that the lack of significant data might be related to the preparation time for tasks, arguing that when students are given preparation time, they might not be able to use this limited period simultaneously for the conceptualization of ideas and the formulation of linguistic forms. However, in this study, even though students were not given a specific preparation time and were allowed to start when they felt ready, there was still no observed impact of creativity on the use of connectives.

Although placement tests are controlled for the MRA results, the extent to which students' proficiency levels solely explain the CAF measurements should also be separately addressed. In the AT, students' proficiency levels significantly accounted for the length of the sentences they constructed, the number of unique words used, the accuracy of their sentences, the fluency and speed of their sentence construction, the reduction of self-repetition or false starts, and the total number of words used.

In the PN task, students' proficiency levels significantly explained the number of subordinate clauses used in their sentences, the greater length of their sentences, the broader textual diversity of their sentences, their ability to construct more accurate sentences, their ability to speak more quickly while constructing sentences with shorter pauses between them, and the greater total number of words produced.

As previously discussed, creativity skills in the AT are only associated with familiarity, total number of words produced, and causal connectives rate among the mentioned CAF measurements. Meanwhile, in the PN task, creativity skills are linked solely to sentence length, end-clause pauses, and students' reduced self-repetition or fewer false starts. Thus, aligning with findings from previous studies (McDonough et al., 2015; Suzuki et al., 2022; Qin, 2024), students' creativity levels affect only a small portion of their speaking performances.

Under the next heading, the relationships between students' preparation times for tasks, CAF measurements, and their creativity skills are examined.

5.1.2. Discussion for the Second Research Question

In order to address the second research question of this study, “In what ways does the task preparation time influence the quality of speech production and creativity,” students were instructed to start when they felt ready without being given a specific preparation time. The preparation durations were recorded in seconds. Subsequently, the relationship between these durations and students’ creativity skills, along with CAF measurements, was examined. The findings indicated that, on average, students spent 33 seconds preparing for the AT, with the shortest preparation time being 4 seconds and the longest 140 seconds. For the PN task, the average preparation time was 41 seconds, with the quickest preparation at 7 seconds and the longest at 158 seconds. This discrepancy between tasks is primarily due to the nature of the PN task, which requires more time for students to examine the images and construct a structure. Additionally, the questions in the AT generally allowed students to draw upon their own experiences, leading to shorter conceptualization times compared to the PN task.

When examining the correlations between students’ preparation times for the speaking tasks and their creativity skills, it was found that there was no significant correlation. Although negative correlations were observed, the hypothesis that preparation times decrease as creativity skills increase could not be substantiated. Consequently, no significant findings were obtained regarding the preparation times of creative students.

In terms of the correlations between preparation times and CAF measurements, it was observed that as preparation time for the AT increased, there was a significant decrease in students’ speech fluency. One major reason for this is that as students spend more time in the conceptualization process, the complexity of the content they structure also increases, leading them to sacrifice fluency in order to articulate more detailed and complex responses. Another possible reason is that students may face difficulties in the conceptualization process when encountering a question about an unfamiliar situation or, even if familiar, may experience anxiety. As Qin (2024) also noted, in such situations, due to cognitive pressure, students allocate their limited attentional resources to L2-specific syntactic structures and lexicon rather than to the fluency of their speaking performance.

Another result observed from the correlations is that as preparation times for the PN task increased, the frequency of causal connectives rate decreased. As previously discussed, students tend to use connectives more frequently when seeking cohesion between different images due to the nature of the task. However, the findings of this study indicate that students focused more on accurately conveying the story rather than on the cause-and-effect relationships between the events depicted in the images. Similar to the results from the AT, students allocated their attentional resources to ensuring an accurate narrative rather than to the connections between events.

In Qin's (2024) study, it was found that when students performed the narration task without any planning, their speaking performance was directly influenced by their creativity. More creative students produced more and more organized ideas, but with a less diversified lexicon. Conversely, students with a five-minute preparation time, while producing more ideas, did so with less fluency. Thus, Qin's (2024) study also demonstrated the negative impact of longer preparation times on students' fluency, which is consistent with the findings of this study.

Examining other studies in the literature on preparation times prior to speaking tasks, Mehnert (1998) observed results by allowing students preparation times of no preparation time, 1 minute, 5 minutes, and 10 minutes. Similarly, O'Grady (2019) provided preparation times of 30 seconds, 1 minute, 5 minutes, and 10 minutes. Gilanlioğlu (2022) investigated the relationship between students' speaking task performance with no planning time and with a 10-minute planning time. Mehnert (1998) found that a one-minute preparation time had the most significant impact on speaking performance, whereas O'Grady (2019) reported that a five-minute preparation time had the most significant impact. Gilanlioğlu (2022) stated that task planning time is necessary but only has positive effects on lexical density, not on lexical complexity.

Therefore, as observed in this study and supported by the literature, a one-minute preparation time is sufficient for task planning. Moreover, it was found that longer planning times do not lead to a more complex and fluent speaking performance.

5.2. Conclusion

To conclude the study, it can be stated that the primary aim was to observe how one of the significant individual differences among students, their creativity, impacts their speaking skills. When examining the relationship between creativity and speaking skills, it appears that the relationship is most strongly observed in the speech conceptualization processes. The effect of creativity on these processes was anticipated to be found by comparatively analyzing students' speaking performances and creativity skills in two tasks with different cognitive demands. The PN task was chosen as the low-demanding task because, due to its predefined content, it was assumed that students would only comment on the events depicted in the images, thus limiting the conceptualization process. The high-demanding task was chosen as the argumentative speaking task, known for its open-ended questions, where students could freely express their ideas, thus demanding more cognitive resources. To measure students' creativity skills, the two main cognitive dimensions of creativity, DT and CT, were assessed using the AUT for DT and the RAT for CT. Additionally, the openness subdimension of the BFI was included as a third criterion for creativity from the social-personality approach.

Participants were selected from the first-year students in the ELT department at Muğla Sıtkı Koçman University through convenience sampling. 72 students completed a personal information form, the RAT, and BFI tests, and then, on the following day, they were individually called to complete the AUT and speaking tasks. The completed tasks were recorded and transcribed into AS-units for analysis, with both pruned and unpruned files prepared for further analysis. CAF measurements and discourse analyses were conducted, followed by correlation analyses using SPSS, and MRA was performed based on significant correlation data.

The results, although generally consistent with the literature, diverged in the use of connectives. Another point of divergence was the students' demonstration of more fluent speaking performance. Unlike the literature, it was noted that students' CT scores showed no correlation with any CAF measurements in either type of task. A potential reason for this could be the length and fatigue induced by the 120-item RAT test. In the second part of the study, it was investigated whether there was any relationship between the students' preparation times for the tasks and their creativity or speaking performance. According to the data, preparation times were not correlated with creativity. In speaking

performance, as preparation times increased, students used fewer connectives in the PN task and articulated their ideas more slowly in the AT. The average preparation times, aligning with previous studies, suggested that an optimal preparation time for reflecting students' speaking performance is around 30 to 45 seconds.

Creativity remains an under-researched and undervalued area in L2 teaching. Reliable and valid studies examining its relationship with speaking are limited. Therefore, this research aims to serve as a clearer and more comprehensive example for future studies. The findings suggest that creativity, recognized as one of the most critical skills of the 21st century, can partly explain speaking performance, particularly in speech conceptualization processes.

5.3. Suggestions

The study emphasizes the importance of creativity as an individual difference and offers a different perspective for future research. Based on the results of this study and previous studies (McDonough et al., 2015; Qin, 2024), it is recommended that speaking tasks be designed considering students' creativity. The study found no significant difference between the speaking performances in high-demanding and low-demanding tasks and their relationship with creativity. Therefore, future research could explore how creativity affects task complexity rather than its role in task planning. Considering previous studies (Mehnert, 1998; O'Grady, 2019) and the findings of this research, a preparation time of 45 seconds to 1 minute is sufficient for optimal speaking performance. Thus, allocating 1 minute for planning time allows students the necessary conceptualization time and enables fluent and accurate speech. Increased preparation times can strain students' cognitive resources, making them unable to meet different demands. Although creativity does not fully explain speaking performance in the argumentative and PN tasks, significant percentages were obtained for key CAF measurements. These results suggest that when making pedagogical decisions, such as task selection and evaluation criteria, students' creativity should not be underestimated, and materials should be developed accordingly.

Despite creativity's undeniable impact, it does not significantly influence speaking performance; proficiency level primarily determines speaking performance,

with creativity contributing moderately. Given the undeniable impact of creativity on language use, creativity courses should be offered at the undergraduate level for students with low creativity, and the importance of creativity should be emphasized in teacher education to help students develop their creative skills. While creativity is an innate personality trait, it can be developed and acquired as a cognitive skill, similar to other cognitive skills. In an era where educational models are becoming more individualized and considering individual differences, extra effort is needed to foster creativity, alongside reducing student anxiety and increasing motivation.

Regarding students' preparation times for tasks, a one-minute preparation time is sufficient for them to demonstrate their speaking performance at full potential, so future studies could limit planning time to one minute. This will not only shorten the duration of the study, benefiting the researcher, but also ensure the study's foundation is based entirely on relevant literature. Another important suggestion for future researchers is to examine creativity as an outcome rather than a predictor variable. Investigating how the creative and playful use of language varies between tasks is a significant gap in the literature. Given the limited exploration of the relationship between creativity and speaking, it is crucial to conduct studies with a larger number of participants, different contexts, and students with different L1 backgrounds to generalize the findings. Examining both the cognitive and social-personality dimensions of students' creativity is essential, and when investigating DT skills, it is important to test not only verbal but also figural creativity to fully observe their creative abilities. Furthermore, future studies could examine originality, elaboration, and flexibility, in addition to fluency, in the context of DT. For studies in Türkiye, considering the Turkish RAT test adapted by Özen et al. (2015) is currently the only available test for Turkish participants, a shorter version with easy, medium, and hard items, with 10 each, totaling 30 items could be used in future research to prevent fatigue from affecting test results.

This study explored the impact of creativity on speaking performance and whether students' preparation times for tasks affect their creativity or speaking performance. There are very few empirical studies on creativity in the international literature, and this study is believed to be pioneering in the Turkish context. The relationship between creativity and spoken proficiency should be further investigated and supported by future research to understand its effects more comprehensively.

REFERENCES

- Abdullah, D. F., Kristina, D., & Sumardi, S. (2020). TOWARD CREATIVITY AND SPEAKING ABILITY OF EFL STUDENTS: A MIXED-METHOD STUDY. *LLT Journal: A Journal on Language and Language Teaching*, 23(2), 207-219. <https://doi.org/10.24071/llt.v23i2.2469>
- Acar, S., & Runco, M. A. (2019). Divergent thinking: New methods, recent research, and extended theory. *Psychology of Aesthetics, Creativity, and the Arts*, 13(2), 153–158. <https://doi.org/10.1037/aca0000231>
- Ahmadi, B., Namaziandost, E., & Etemadfar, P. (2020). A Comparative Study of Using Analytic and Holistic Methods of Scoring in Measuring Speaking Skill Among Iranian Second-grade University EFL Students. *Journal of English Language Studies*, 5(1), 15-32. <https://doi.org/10.30870/jels.v5i1.6397>
- Albert, Á. (2006). Learner creativity as a potentially important individual variable: Examining the relationships between learner creativity, language aptitude and level of proficiency. In M. Nikolov & J. Horváth (Eds.), *Empirical studies in English applied linguistics* (pp. 77-98). Pécs: Lingua Franca Csopor.
- Albert, A. (2012). Creativity. In Robinson, P. (Ed.), *Routledge encyclopedia of second language acquisition* (pp. 144-146). London: Routledge.
- Albert, Á., & Kormos, J. (2011). Creativity and Narrative Task Performance: An Exploratory Study. *Language Learning*, 61(s1), 73–99. <https://doi.org/10.1111/j.1467-9922.2011.00643.x>
- Al-Sibai, D. (2004). *Promoting Oral fluency of Second Language Learners: Educational Linguistics*. Department of English. King Saud University.
- Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology*, 43(5), 997–1013. <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), 357–376. <https://doi.org/10.1037/0022-3514.45.2.357>
- Amabile, T. M., & Gryskiewicz, N. D. (1989). The creative environment scales: Work environment inventory. *Creativity Research Journal*, 2(4), 231–253. <https://doi.org/10.1080/10400418909534321>
- Amoah, S., & Yeboah, J. (2021). The speaking difficulties of Chinese EFL learners and their motivation towards speaking the English language. *Journal of Language and Linguistic Studies*, 17(1), 56-69.
- Ananiadou, K., & Claro, M. (2009). 21st Century Skills and Competences for New Millennium Learners in OECD Countries. *OECD Education Working Papers*, 41. Paris: OECD Publishing. <https://doi.org/10.1787/218525261154>
- Anbe, K. (2016). *Three Studies on Culture and Creativity: The Influence of Cultural Syndromes, Multicultural Personality, and Multicultural Experience on Divergent, Convergent, and Holistic Tasks of Creativity*. [Doctoral dissertation] Honolulu, University of Hawaii at Manoa.
- Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: development and validation of the team climate inventory. *Journal of Organizational Behavior*, 19(3), 235–258. [https://doi.org/10.1002/\(SICI\)1099-1379\(199805\)19:3<235::AID-JOB837>3.0.CO;2-C](https://doi.org/10.1002/(SICI)1099-1379(199805)19:3<235::AID-JOB837>3.0.CO;2-C)
- Apuke, O.D., (2017). Quantitative research methods: A synopsis approach. *Kuwait Chapter of Arabian Journal of Business and Management Review*, 33(5471), 1-8.
- Arshavskaya, E. (2015). Creative writing assignments in a second language course: A way to engage less motivated students. *InSight: A Journal of Scholarly Teaching*, 10, 68-78. <https://doi.org/10.46504/10201506ar>
- Asher, J. (1966). *The learning strategy of the total physical response: A review*. *Modern Language Journal* 50: 79–84.
- Baas, D., Castelijns, J., Vermeulen, M., Martens, R., & Segers, M. (2015). The relation between Assessment for Learning and elementary students' cognitive and

- metacognitive strategy use. *The British journal of educational psychology*, 85(1), 33–46. <https://doi.org/10.1111/bjep.12058>
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing*, 17(1), 1-42.
<http://dx.doi.org/10.1191/026553200675041464>
- Baddeley, A. (2003). Working Memory: Looking Back and Looking Forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Bailey, K. M. (2005). *Practical English language teaching: speaking*. McGraw-Hill ESL/ELT.
- Bailey, K. M. (2020). *Teaching Listening and Speaking in second and Foreign Language Context*. London: Great Britain.
- Barbot, B., Besançon, M., & Lubart, T. I. (2011). Assessing creativity in the classroom. *The Open Education Journal*, 4, 124-132.
<https://doi.org/10.2174/1874920801104010058>
- Baron, N. (2000). *Alphabet to Email*. London: Routledge.
- Batey, M. (2012). The measurement of creativity: From definitional consensus to the introduction of a new heuristic framework. *Creativity Research Journal*, 24(1), 55–65. <https://doi.org/10.1080/10400419.2012.649181>
- Beaman, K. (1984). Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In D. Tannen (Ed.), *Coherence in Spoken and Written Discourse* (pp. *Language, Exploring Orality and Literacy*, 45–80). Ablex Publishing Corporation.
- Beaty, R. E., Benedek, M., Wilkins, R. W., Jauk, E., Fink, A., Silvia, P. J., Hodges, D. A., Koschutnig, K., & Neubauer, A. C. (2014a). Creativity and the default network: A functional connectivity analysis of the creative brain at rest. *Neuropsychologia*, 64, 92–98.
<https://doi.org/10.1016/j.neuropsychologia.2014.09.019>

- Beaty, R. E., Nusbaum, E. C., & Silvia, P. J. (2014b). Does insight problem solving predict real-world creativity? *Psychology of Aesthetics, Creativity, and the Arts*, 8(3), 287–292. <https://doi.org/10.1037/a0035727>
- Besemer, S. P., & O'Quin, K. (1999). Confirming the three-factor creative product analysis matrix model in an American sample. *Creativity Research Journal*, 12(4), 287–296. https://doi.org/10.1207/s15326934crj1204_6
- Bezerra, R. L. M., Alves, R. J. R., & Azoni, C. A. S. (2022). Creativity and its relationship with intelligence and reading skills in children: an exploratory study. *Psicologia: Reflexao e Critica*, 35, 17. <https://doi.org/10.1186/s41155-022-00221-3>
- Bhimavarapu, V. (2023). The Impact of Generative AI on Human Productivity in Creative Writing. *Journal of Student Research*, 12(3). <https://doi.org/10.47611/jsrhs.v12i3.4780>
- Bialystok, E., Craik, F. I. M., & Luk, G. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, 16(4), 240–250. <https://doi.org/10.1016/j.tics.2012.03.001>
- Biber D., S. Conrad, & G. Leech. (2002). *Longman Student Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Blicblau, A. S., & Steiner, J. M. (1998). Fostering Creativity Through Engineering Projects. *European Journal of Engineering Education*, 23(1), 55–65. <https://doi.org/10.1080/0304379980230107>
- Bloland, P. A. (1987). Leisure as a campus resource for fostering student development. *Journal of Counseling & Development*, 65(6), 291–294. <https://doi.org/10.1002/j.1556-6676.1987.tb01288.x>
- Bloomfield, J., & Fisher, M. J. (2019). Quantitative research design. *Journal of the Australasian Rehabilitation Nurses Association*, 22(2), 27-30. <https://search.informit.org/doi/10.3316/informit.738299924514584>
- Boersma, P. & Weenink, D. (2024). *Praat: doing phonetics by computer* [Computer software]. Version 6.4.13, retrieved 10 June 2024 from <http://www.praat.org/>

- Bower, J., Acar, S., & Kurşuncu, U. (2023). Measuring creativity in academic writing: An analysis of essays in Advanced Placement Language and Composition. *Journal of Advanced Academics*. <https://doi.org/10.1177/1932202X231211633>
- Brophy, D. R. (2001). Comparing the Attributes, Activities, and Performance of Divergent, Convergent, and Combination Thinkers. *Creativity Research Journal*, 13(3–4), 439–455. https://doi.org/10.1207/S15326934CRJ1334_20
- Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge University Press.
- Brown, H. D., & Abeywickrama, P. (2004). *Language assessment*. Principles and Classroom Practices. White Plains, NY: Pearson Education.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy*. Pearson.
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 23-46. Amsterdam/Philadelphia: John Benjamins.
- Burns, A. (2017). Research and the teaching of speaking in the second language classroom. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning, Volume III*. New York: Routledge.
- Burns, A., & Joyce, H. (1997). *Focus on Speaking*. National Centre for English Language Teaching and Research, Macquarie University, Sydney, New South Wales, Australia 2109.
- Burns, N., Grove, S. K., & Gray, J. (2015). *Understanding nursing research: building an evidence-based practice*. 6th edition. St. Louis, Missouri, Elsevier.
- Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review*, 67, 380–400. <http://dx.doi.org/10.1037/h0040373>
- Campbell, M. (2022). *The effects of noticing training, model input and task repetition on L2 speech production* [Doctoral dissertation] Macquarie University.

- Carlson, S.M. & Meltzoff, A.N. (2008). Bilingual Experience and Executive Functioning in Young Children. *Developmental Science*, 11, 282-298.
<http://dx.doi.org/10.1111/j.1467-7687.2008.00675.x>
- Celce-Murcia, M., Brinton, D., & Snow, A. (2014). *Teaching English as a Second or Foreign Language* (4th ed.). Boston, MA: Heinle ELT.
- Cemalcilar, Z., Baruh, L., Kezer, M., Soto, C. J., Sumer, N., & John, O. P. (2021). Testing the BFI-2 in a non-WEIRD community sample. *Personality and Individual Differences*, 182, 111087. <https://doi.org/10.1016/j.paid.2021.111087>
- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535-544.
[https://doi.org/10.1016/S0346-251X\(97\)00046-8](https://doi.org/10.1016/S0346-251X(97)00046-8)
- Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly Openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences*, 19(4), 524–529.
<https://doi.org/10.1016/j.lindif.2009.06.004>
- 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
- Couger, J. D., Higgins, L. F., & McIntyre, S. C. (1993). (Un)structured creativity in information systems organizations. *MIS Quarterly*, 17, 375-397.
<https://doi.org/10.2307/249584>
- Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion Volume*. Council of Europe Publishing.
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.
- Curran, C. (1976) *Counseling-Learning in Second Languages*. Illinois: Apple River Press.
- Cushen, P. J., & Wiley, J. (2011). Aha! Voila! Eureka! Bilingualism and insightful problem solving. *Learning and Individual Differences*, 21(4), 458–462.
<https://doi.org/10.1016/j.lindif.2011.02.007>

- Dacey, J. (1999). Concepts of creativity: A history. In M. A. Runco and S. R. Pritzker (Eds.), *Encyclopedia of creativity*. CA: Academic Press.
- Davies, C., & Fisher, M. (2018). Understanding research paradigms. *Journal of the Australasian Rehabilitation Nurses Association*, 21(3), 21–25.
<https://search.informit.org/doi/10.3316/informit.160174725752074>
- Davies, D., Jindal-Snape, D., Collier, C., Digby, R., Hay, P., & Howe, A. (2013). Creative learning environments in education-A systematic literature review. *Thinking Skills and Creativity*, 8, 80-91.
<https://doi.org/10.1016/j.tsc.2012.07.004>
- Dawson, P. (2005). *Creative writing and the new humanities*. London: Routledge.
- de Chantal, P. L., & Markovits, H. (2017). The capacity to generate alternative ideas is more important than inhibition for logical reasoning in preschool-age children. *Memory & cognition*, 45, 208-220. <https://doi.org/10.3758/s13421-016-0653-4>
- de Jong, N. H. (2023). Assessing second language speaking proficiency. *Annual Review of Linguistics*, 9(1), 541-560. <https://doi.org/10.1146/annurev-linguistics-030521-052114>
- de Jong, N. H., Groenhout, R., Schoonen, R., & Hulstijn, J. H. (2015). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 36, 223-243.
<https://doi.org/10.1017/S0142716413000210>
- deVink, I. C., Willemsen, R. H., Lazonder, A. W., & Kroesbergen, E. H. (2022). Creativity in mathematics performance: The role of divergent and convergent thinking. *British Journal of Educational Psychology*, 92(2), Article e12459.
<https://doi.org/10.1111/bjep.12459>
- Dewaele, J.-M. (2012). “Personality. Personality traits as independent and dependent variables”. In *Psychology For Language Learning: Insights from Research, Theory and Practice*, S. Mercer, S. Ryan and M. Williams (eds.), 42–58. Basingstoke: Palgrave Macmillan.
- Dewaele, J.-M., & Furnham, A. (2000). Personality and speech production: a pilot study of second language learners. *Personality and Individual Differences*, 28(2), 355–365. [https://doi.org/10.1016/S0191-8869\(99\)00106-3](https://doi.org/10.1016/S0191-8869(99)00106-3)

- Dil, Y. (2009). EFL Learners' communication obstacles. *Electronic Journal of Social Sciences*, 8(29), 84–100.
- Domino, G. (1994). Assessment of creativity with the ACL: An empirical comparison of four scales. *Creativity Research Journal*, 7(1), 21–33.
<https://doi.org/10.1080/10400419409534506>
- Dörnyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research : LTR*, 4(3), 275–300.
<https://doi.org/10.1177/136216880000400305>
- D'Souza, R. (2021). What characterises creativity in narrative writing, and how do we assess it? Research findings from a systematic literature search. *Thinking skills and creativity*, 42, 100949. <https://doi.org/10.1016/j.tsc.2021.100949>
- Duff, P. A. (2014). Case study research on language learning and use. *Annual Review of Applied Linguistics*, 34, 233-255. <https://doi.org/10.1017/S0267190514000051>
- Elia, F., & Irianti, D. (2015). Psychological factors that hinder students from speaking in English. *Journal of English for Academic*, 2(2), 50-61.
- ETS. (2012). The official guide to the TOEFL® test. New York, NY :McGraw-Hill.
- Fasko, D. (2001). Education and Creativity. *Creativity Research Journal*, 13, 317-327.
https://doi.org/10.1207/S15326934CRJ1334_09
- Fellbaum, C. (1988). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Fernández-Fontecha, A. (2021). The role of learner creativity in L2 semantic fluency. An exploratory study. *System (Linköping)*, 103, 102658-.
<https://doi.org/10.1016/j.system.2021.102658>
- Fishkin, A. S., & Johnson, A. S. (1998). Who is creative? Identifying children's creative abilities. *Roeper Review*, 21(1), 40-46.
<https://doi.org/10.1080/02783199809553925>
- Florez, M. C. (1999). Improving adult English language learners' speaking skills. *The CAELA Guide For Adult ESL Trainers*, 16-20.
- Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36, 98-116. <https://doi.org/10.1017/S0267190515000082>

- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, 21(3), 354-375.
<https://doi.org/10.1093/applin/21.3.354>
- Fulcher, G. (2014). *Testing Second Language Speaking* (1st ed.). Routledge.
<https://doi.org/10.4324/9781315837376>
- Fulcher, Glenn. (2003). *Testing second language speaking*. London: Pearson.
- Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences - PNAS*, 112(33), 10336–10341.
<https://doi.org/10.1073/pnas.1502134112>
- Gan, Z. (2012). Understanding L2 speaking problems : implications for ESL curriculum development in a teacher training institution in Hong Kong. *The Australian Journal of Teacher Education*, 37(1), 43–59.
<https://doi.org/10.14221/ajte.2012v37n1.4>
- Gardner, R. C., & MacIntyre, P. D. (1992). A student's contributions to second language learning. Part I: Cognitive variables. *Language Teaching*, 25(4), 211–220. <https://doi.org/10.1017/S026144480000700X>
- Gebhard, G.J. (2000). *Teaching English as a Foreign or Second Language*. USA: The University of Michigan Press.
- Gilanlioğlu, I. (2022). Influence of planning time and task type on lexis in L2 oral performance. *Bartın University Journal of Faculty of Education*, 11(3), 559-575.
- Glăveanu, V. P. (2013). Creativity and Folk Art: A Study of Creative Action in Traditional Craft. *Psychology of Aesthetics, Creativity, and the Arts*, 7(2), 140–154. <https://doi.org/10.1037/a0029318>
- Glăveanu, V. P. (2017). Introducing Creativity and Culture, the Emerging Field. In *The Palgrave Handbook of Creativity and Culture Research* (pp. 1–12). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-46344-9_1
- Goh, C.C.M. (2016). Teaching Speaking. In: Renandya, W., Widodo, H. (eds) *English Language Teaching Today*. English Language Education, vol 5. Springer, Cham. https://doi.org/10.1007/978-3-319-38834-2_11
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26–42. <http://dx.doi.org/10.1037/1040-3590.4.1.26>

- Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology*, 37, 1398-1405.
- Göçen, G. (2019). The effect of creative writing activities on elementary school students' creative writing achievement, writing attitude and motivation. *Journal of Language and Linguistic Studies*, 15(3), 1032-1044.
<https://doi.org/10.17263/jlls.631547>
- Guilford, J. P. (1950). Creativity. *American Psychologist*, 5(9), 444-454.
<https://doi.org/10.1037/h0063487>
- Guilford, J. P. (1967a). Creativity: Yesterday, today and tomorrow. *Journal of Creative Behavior*, 1(1), 3-14. <https://doi.org/10.1002/j.2162-6057.1967.tb00002.x>
- Guilford, J. P. (1967b). The nature of human intelligence. New York: McGraw-Hill
- Guilford, J.P. (1959) *Traits of Creativity*. In: Anderson, H.H., Ed., Creativity and Its Cultivation, Harper & Row, New York, 142-161.
- Haerazi, H., Irawan, L. A., Suadiyatno, T., & Hidayatullah, H. (2020). Triggering Preservice Teachers' Writing Skills through Genre-Based Instructional Model Viewed from Creativity. *International Journal of Evaluation and Research in Education*, 9(1), 234-244. <https://doi.org/10.11591/ijere.v9i1.20395>
- Halliday, M. A. K. (1979). *Differences between spoken and written language: Some implications for literacy teaching*. In G. Page, J. Elkins and B. O'Connor (eds) Communication through reading: Proceedings of the Fourth Australian Reading Conference Vol. 2. Adelaide: Australian Reading Association: 37-52.
- Halliday, M. A. K. (1989). *Spoken and Written Language*. Oxford: Oxford University Press.
- Hammerly, H. (1990). *Fluency and accuracy: Toward balance in language teaching and learning*. Clevedon: Multilingual Matters.
- Hansen, L. B., Macizo, P., Duñabeitia, J. A., Saldaña, D., Carreiras, M., Fuentes, L. J., & Bajo, M. T. (2016). Emergent bilingualism and working memory development in school aged children. *Language Learning*, 66(S2), 51-75.
<https://doi.org/10.1111/lang.12170>
- Hatipoğlu, Ç. (2021). Testing and assessment of speaking skills, test task types, and sample test items. In Servet Çelik, Handan Çelik and Christine Coombe (Eds.),

Language assessment and test preparation in English as a foreign language (EFL) education. 119-173. Ankara: Vizetek.

- Hatipoğlu, Ç. (2021). Testing and assessment of speaking skills, test task types, and sample test items. *Language assessment and test preparation in English as a foreign language (EFL) education*, 119-173.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- Hennessey, B. A., & Amabile, T. M. (2010). Creativity. *Annual Review of Psychology*, 61, 569–598. <https://doi.org/10.1146/annurev.psych.093008.100416>
- Hinkel, E. (2006). Current Perspectives on Teaching the Four Skills. *TESOL Quarterly*, 40(1), 109–131. <https://doi.org/10.2307/40264513>
- Hoelcher, M., & Schubert, M. (2015). Potential and problems of existing creativity and in novation indices. *Creativity Research Journal*, 27, 1-15. <https://doi.org/10.1080/10400419.2015.992656>
- Horn, D., & Salvendy, G. (2006). Consumer-based assessment of product creativity: A review and reappraisal. *Human Factors and Ergonomics in Manufacturing*, 16(2), 155–175. <https://doi.org/10.1002/hfm.20047>
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal*, 70, 125-132. <http://dx.doi.org/10.1111/j.1540-4781.1986.tb05256.x>
- Horzum, M. B., Ayas, T., & Padır, M. A. (2017). Beş faktör kişilik ölçeğinin Türk kültürüne uyarlanması adaptation of big five personality traits scale to Turkish culture. *Sakarya University Journal of Education*, 7(2), 398-408. <http://dx.doi.org/10.19126/suje.298430>
- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency. *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*, 32, 1-20. <https://doi.org/10.1075/llt.32.01hou>
- Hu, W., & Adey, P. (2002). A scientific creativity test for secondary school students. *International Journal of Science Education*, 24(4), 389–403. <https://doi.org/10.1080/09500690110098912>

- Hymes, D. (1972). On communicative competence, 269-293. J. B. Pride, & J. Holmes (Eds.). *Sociolinguistics*, 269–293.
- Innamorato, G. (1998). Creativity in the development of scientific giftedness: Educational implications. *Roeper Review*, 21(1), 54–59.
<https://doi.org/10.1080/02783199809553932>
- Inoue, C. (2016). A comparative study of the variables used to measure syntactic complexity and accuracy in task-based research. *Language Learning Journal*, 44(4), 1-19. <https://doi.org/10.1080/09571736.2015.1130079>
- Inoue, C. (2016). A comparative study of the variables used to measure syntactic complexity and accuracy in task-based research. *The Language Learning Journal*, 44(4), 487–505. <https://doi.org/10.1080/09571736.2015.1130079>
- Ippolito, D., Yuan, A., Coenen, A., & Burnam, S. (2022). Creative writing with an ai-powered writing assistant: perspectives from professional writers. *ArXiv*, 1-18.
<https://doi.org/10.48550/arXiv.2211.05030>
- Isaacs, T. (2016). Assessing speaking. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment*, 131-146. Berlin: DeGruyter Mouton.
- Isaksen, S. G., Lauer, K. J., & Ekvall, G. (1999). Situational Outlook Questionnaire: A measure of the climate for creativity and change. *Psychological Reports*, 85(2), 665–674. <https://doi.org/10.2466/PR0.85.6.665-674>
- Ivcevic, Z., & Hoffmann, J. (2019). Emotions and creativity: From process to person and product. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (2nd ed., pp. 273–295). Cambridge University Press.
<https://doi.org/10.1017/9781316979839.015>
- Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning*, 63(s1), 87–106. <https://doi.org/10.1111/j.1467-9922.2012.00739.x>
- Javaid, S. F., & Pandarakalam, J. P. (2021). The association of creativity with divergent and convergent thinking. *Psychiatria Danubina*, 33(2), 133–139.
- Jiang, Y., & Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System (Linköping)*, 82, 13–25.
<https://doi.org/10.1016/j.system.2019.02.017>

- John, O. P. (1989). Towards a taxonomy of personality descriptors. In D. M. Buss & N. Cantor (Eds.), *Personality psychology: Recent trends and emerging directions* (pp. 261–271). New York, NY: SpringerVerlag. http://dx.doi.org/10.1007/978-1-4684-0634-4_20
- John, O. P. (1990). The “Big Five” factor taxonomy: Dimensions of personality in the natural language and questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66 –100). New York, NY: Guilford Press.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory – versions 4a and 5*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- Johnson, K. (1996). The role of theory in L2 teacher education. *TESOL Quarterly*, 30, 765-771.
- Johnson, V. E. (1985). The place of general achievement testing in the Freshman English language program. *Collection of Humanities*, 32, 28-54. Accessed at [https://m-repo.lib.meiji.ac.jp/dspace/bitstream/10291/11979/1/jinbunkagakuronshu_32_\(35\).pdf](https://m-repo.lib.meiji.ac.jp/dspace/bitstream/10291/11979/1/jinbunkagakuronshu_32_(35).pdf)
- Karimpour, S., & Chopoglou, M. A. M. (2014). The relationship between creativity and Iranian EFL learners’ speaking skill. *Journal of Educational and Management Studies*, 4(4), 877-888.
- Karwowski, M., & Lebuda, I. (2016). The big five, the huge two, and creative self-beliefs: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 10(2), 214–232. <https://doi.org/10.1037/aca0000035>
- Karya, P. J., Takarroucht, K., Zano, K., & Zamorano, A. (2022). Developing the prototype of picture-based learning materials in the teaching of speaking skills. *Journal of Language and Literature Studies*, 2(2), 109-116. <http://dx.doi.org/10.36312/jolls.v2i2.763>
- Kaufman, J. C., & Baer, J. (2004). Hawking's Haiku, Madonna's math: Why it is hard to be creative in every room of the house. In R. J. Sternberg, E. L. Grigorenko, & J. L. Singer (Eds.), *Creativity: From potential to realization* (pp. 3-19). Washington, DC: American Psychological Association.

- Kaufman, J. C., & Baer, J. (2012). Beyond new and appropriate: Who decides what is creative? *Creativity Research Journal*, 24, 83-91.
<https://doi.org/10.1080/10400419.2012.649237>
- Kaufman, J. C., & Glăveanu, V. P. (2019). A review of creativity theories: What questions are we trying to answer? In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (2nd ed., pp. 27–43). Cambridge University Press. <https://doi.org/10.1017/9781316979839.004>
- Kaufman, J. C., & Sternberg, R. J. (2010). *Cambridge Handbook of Creativity* (1st ed.). Cambridge University Press.
- Kaufman, J. C., Evans, M. L., & Baer, J. (2010). The American Idol Effect: Are students good judges of their creativity across domains? *Empirical Studies of the Arts*, 28, 3-17. <https://doi.org/10.2190/EM.28.1.b>
- Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, consistency, creativity, and the consensual assessment technique: New evidence of reliability. *Thinking Skills and Creativity*, 2(2), 96–106. <https://doi.org/10.1016/j.tsc.2007.04.002>
- Kaufman, J. C., Plucker, J. A., & Baer, J. (2008). *Essentials of creativity assessment*. New York: Wiley.
- Kaufman, S. B. (2013). Opening up openness to experience: A four-factor model and relations to creative achievement in the arts and sciences. *The Journal of Creative Behavior*, 47(4), 233–255. <https://doi.org/10.1002/jocb.33>
- Kaufman, S. B., Quilty, L. C., Grazioplene, R. G., Hirsh, J. B., Gray, J. R., Peterson, J. B., & DeYoung, C. G. (2016). Openness to Experience and Intellect Differentially Predict Creative Achievement in the Arts and Sciences. *Journal of personality*, 84(2), 248–258. <https://doi.org/10.1111/jopy.12156>
- Kaulfers, Walter V. (1944). Wartime development in modern-language achievement testing. *Modern Language Journal*, 28(2). 136-150.
<http://dx.doi.org/10.1111/j.1540-4781.1944.tb04835.x>
- Keyes, D. (2004). *Flowers for Algernon* (1st Harvest ed.). Harcourt.
- Kharkhurin, A. V. (2017). Language mediated concept activation in bilingual memory facilitates cognitive flexibility. *Frontiers in Psychology*, 8, 1067–1067.
<https://doi.org/10.3389/fpsyg.2017.01067>

- Kitano, K. (2001). Anxiety in the College Japanese Language Classroom. *The Modern Language Journal* (Boulder, Colo.), 85(4), 549–566.
<https://doi.org/10.1111/0026-7902.00125>
- Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum Associates.
- Krashen, S. (1987). *Principles and Practice in Second Language Acquisition*. Herts: Prentice-Hall International(UK).
- Kudrowitz, B., & Diplo, C. (2013). When Does a Paper Clip Become a Sundial? Exploring the Progression of Originality in the Alternative Uses Test. *Journal of Integrated Design & Process Science*, 17(4), 3–18. <https://doi.org/10.3233/jid-2013-0018>
- Kuhn, J. T., & Holling, H. (2009). Exploring the nature of divergent thinking: A multilevel analysis. *Thinking Skills and Creativity*, 4(2), 116-123.
<https://doi.org/10.1016/j.tsc.2009.06.004>
- Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of second language writing*, 17(1), 48-60. <https://doi.org/10.1016/j.jslw.2007.08.003>
- Kuiken, F., & Vedder, I. (2014). Rating written performance: What do raters do and why? *Language Testing*, 31(3), 329-348.
<https://doi.org/10.1177/0265532214526174>
- Kumar, L., & Kumari, P. (2016). A Study of Convergent Thinking and Divergent Thinking among Secondary School Students in relation to Ethnicity, Locale, Types of Institutions and Sex. *Editor's Note 3 Culturally Designed Pedagogy in India 5*, 41(4), 111.
- Kumaravadivelu, B. (2006). TESOL methods: Changing tracks, challenging trends. *TESOL Quarterly*, 40(1), 59–81. <https://doi.org/10.2307/40264511>
- Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics*, 35(5), 60-61.
<https://doi.org/10.1093/applin/amu047>

- Lambert, C., Kormos, J., & Minn, D. (2017). TASK REPETITION AND SECOND LANGUAGE SPEECH PROCESSING. *Studies in Second Language Acquisition*, 39(1), 167–196. <https://doi.org/10.1017/S0272263116000085>
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics*, 27(4), 590–619. <https://doi.org/10.1093/applin/aml029>
- Larsen-Freeman, D. (2011). A complexity theory approach to second language development/acquisition. In *Alternative approaches to second language acquisition* (pp. 48-72). Routledge.
- Larson-Hall, J. (2010). *A Guide to Doing Statistics in Second Language Research Using SPSS*. New York: Routledge.
- Lazaraton, A. (2014). Second language speaking. In M. Celce-Murcia, D. M. Brinton & M. A. Snow (Eds.), *Teaching English as a second or foreign language* (pp. 106–120). Boston: National Geographic Learning/Heinle Cengage Learning.
- Leahy, A., Cantrell, M. & Swander, M. (2014). 1 Theories of Creativity and Creative Writing Pedagogy. In *The Handbook of Creative Writing* (pp. 11-23). Edinburgh: Edinburgh University Press.
<https://doi.org/10.1515/9780748689774-004>
- Lee, C. S., Huggins, A. C., & Therriault, D. J. (2014). A measure of creativity or intelligence? Examining internal and external structure validity evidence of the remote associates test. *Psychol. Aesth. Creat. Arts*, 8, 446–460.
<https://dx.doi.org/10.1037/a0036773>
- Leech, G. M. Deuchar & R. Hoogenraad. (1982). *English Grammar for Today*. Basingstoke: Macmillan.
- Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggensbach (Ed.). *Perspectives on fluency*, 25-42. Ann Arbor, MI: The University of Michigan Press.
- Lindauer, M. S., & Long, D. A. (1986). The criteria used to judge art: Marketplace and academic comparisons. *Empirical Studies of the Arts*, 4, 163-174.

- Lintunen, P., & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: comparing the incomparable?. *Research in Language*, 12(4), 377-399. <https://doi.org/10.1515/rela-2015-0005>
- Lynch, M. L. (2015). *Assessing creativity via divergent thinking in residential camp settings*. [Unpublished doctoral dissertation] University of New Hampshire.
- Maghfiroh, T. F. D., & Hikmat, M. H. (2023). The Students' Perspectives About the Effectiveness Argumentative Speaking Tasks Given on English Class. *Jurnal Educatio FKIP UNMA*, 9(2), 541-547. <https://doi.org/10.31949/educatio.v9i2.4843>
- Malvey, H. (2017). *Persuasive language in high school students: differences in syntactic complexity in spoken and written language* [Unpublished master's thesis]. Texas Christian University.
- Marshall, S. (1974). *Creative writing*. London: Macmillan.
- Mason, J. H. (2003). *The value of creativity : the origins and emergence of a modern belief*. Ashgate.
- Mathisen, G. E., & Einarsen, S. (2004). A Review of Instruments Assessing Creative and Innovative Environments Within Organizations. *Creativity Research Journal*, 16(1), 119–140. https://doi.org/10.1207/s15326934crj1601_12
- McCarthy, P. M., & Jarvis, S. (2010). MTLT, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392. <https://doi.org/10.3758/BRM.42.2.381>
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, 52(6), 1258–1265. <https://doi.org/10.1037/0022-3514.52.6.1258>
- McCrae, R. R., & Costa, P. T., Jr. (2008). The Five-Factor Theory of personality. In O. P. John, R. W. Robins & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 159–181). New York, NY: Guilford.
- 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. <https://doi.org/10.1002/tesq.211>

- McInnis, L. (2009). *Analyzing English L1 and L2 paraphrasing strategies through concurrent verbal report and stimulated recall protocols*. [Unpublished MA Thesis]. University of Toronto.
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.
- Mednick, S. A. (1962). The associative basis of the creative process. *Psychol. Rev.*, 44, 220–232. <https://dx.doi.org/10.1037/h0048850>
- Mehnert, U. (1998). THE EFFECTS OF DIFFERENT LENGTHS OF TIME FOR PLANNING ON SECOND LANGUAGE PERFORMANCE. *Studies in Second Language Acquisition*, 20(1), 83–108. <https://doi.org/10.1017/S0272263198001041>
- Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. In *The Routledge handbook of instructed second language acquisition*. 50-68. Routledge.
- Milićević, J., & Tsedryk, A. (2011). Assessing and improving paraphrasing competence in FSL. In *Proceedings of the 5th international conference on meaning-text theory* (pp. 175-184).
- Misbah, N. H., Mohamad, M., Md Yunus, M., & Ya'acob, A. (2017). Identifying the Factors Contributing to Students' Difficulties in the English Language Learning. *Creative Education*, 8, 1999-2008. <https://doi.org/10.4236/ce.2017.813136>
- Moos, R. (1986). *Work Environment Scale: Manual*. Palo Alto, CA: Consulting Psychologists Press.
- Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of experimental child psychology*, 114(2), 187–202. <https://doi.org/10.1016/j.jecp.2012.09.002>
- Mourgues, C. V., Preiss, D. D., & Grigorenko, E. L. (2014). Reading skills, creativity, and insight: Exploring the connections. *The Spanish Journal of Psychology*, 17, 1-10. <https://doi.org/10.1017/sjp.2014.59>
- Munandar, U. (2021). *Pengembangan kreativitas anak berbakat*. PT Rineka Cipta.

- Nakhalah, D.A. (2016). Problems and Difficulties of Speaking That Encounter English Language Students at Al Quds Open University.
- Nan, C. (2018). Implications of interrelationship among four language skills for high school English teaching. *Journal of Language Teaching and Research*, 9(2), 418-423. <http://dx.doi.org/10.17507/jltr.0902.26>
- Nicolau, M. F. S., & Sukamto, K. E. (2018). Written vs Spoken Narratives by Indonesian ESL Young Learners: A Case Study. *Journal of Language and Literature*, 18(2), 125-136. <https://doi.org/10.24071/joll.2018.180203>
- Niu, W., & Sternberg, R. J. (2001). Cultural influences on artistic creativity and its evaluation. *International Journal of Psychology*, 36(4), 225–241. <https://doi.org/10.1080/00207590143000036>
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, 30(4), 555-578. <https://doi.org/10.1093/applin/amp044>
- Norris, J.M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555-578. <https://doi.org/10.1093/applin/amp044>
- Nosratinia, M., & Razavi, F. (2016). Writing complexity, accuracy, and fluency among EFL learners: Inspecting their interaction with learners' degree of creativity. *Theory and practice in language studies*, 6(5), 1043. <https://doi.org/10.17507/tpsls.0605.19>
- O'Grady, S. (2019). The impact of pre-task planning on speaking test performance for English-medium university admission. *Language Testing*, 36(4), 505–526. <https://doi.org/10.1177/0265532219826604>
- O'Neil, H. F., Abedi, J., & Spielberger, C. D. (2012). The measurement and teaching of creativity. In *Motivation: Theory and research* (pp. 245-263). Routledge.
- Otto, I. (1998). The Relationship Between Individual Differences in Learner Creativity and Language Learning Success. *TESOL Quarterly*, 32(4), 763–773. <https://doi.org/10.2307/3588011>
- Özen, G., Doğan, A. & Cinan, S. (2015). Uzak Bağlantılar Testi: Norm ve güvenirlik çalışması. *Psikoloji Çalışmaları Dergisi*, 35(2), 25-46.

- Pakula, H.-M. (2019). Teaching speaking. *Apples (Jyväskylä, Finland)*, 13(1), 95–111.
<https://doi.org/10.17011/apples/urn.201903011691>
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590-601. <https://doi.org/10.1093/applin/amp045>
- Pallotti, G. (2015). A simple view of linguistic complexity. *Second Language Research*, 31(1), 117-134. <https://doi.org/10.1177/0267658314536435>
- Park, S. (2022). A Corpus-Based Comparison of Syntactic Complexity in Spoken and Written Learner Language. *Canadian Journal of Applied Linguistics / Revue canadienne de linguistique appliquée*, 25(2), 47-70.
<https://doi.org/10.37213/cjal.2022.32477>
- Pawlak, M. (2016). Assessment of Language Learners' Spoken Texts: Overview of Key Issues. In: Chodkiewicz, H., Steinbrich, P., Krzemińska-Adamek, M. (eds) *Working with Text and Around Text in Foreign Language Environments*. Second Language Learning and Teaching. Springer, Cham. https://doi.org/10.1007/978-3-319-33272-7_6
- Pietilä, P. (1999). *L2 Speech: Oral Proficiency of Students of English at University Level*. Anglicana Turkuensia No 19. Turku: University of Turku.
- Piffer, D. (2012). Can creativity be measured? An attempt to clarify the notion of creativity and general directions for future research. *Thinking Skills and Creativity*, 7(3), 258–264. <https://doi.org/10.1016/j.tsc.2012.04.009>
- Pinker, S. (2007). *The language instinct : how the mind creates language* (1st Harper Perennial Modern Classics ed.). HarperPerennial ModernClassics.
- Pishghadam, R., & Mehr, F. J. (2011). Learner Creativity and Performance in Written Narrative Tasks. *World Journal of Education*, 1(2), 115-125.
<https://doi.org/10.5430/wje.v1n2p115>
- Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist*, 39(2), 83–96.
https://doi.org/10.1207/s15326985ep3902_1
- Plucker, J. A., Makel, M. C., & Qian, M. (2019). Assessment of creativity. In J. C. Kaufman & R. J. Sternberg (Eds.), *The Cambridge handbook of creativity* (2nd

- ed., pp. 44–68). Cambridge University Press.
<https://doi.org/10.1017/9781316979839.005>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912.
<https://doi.org/10.1111/lang.12079>
- Popov, A. (1992). Creativity and reading comprehension. *The Journal of Creative Behavior*, 26(3), 206–212. <https://doi.org/10.1002/j.2162-6057.1992.tb01177.x>
- Priest, T. (2006). Self-evaluation, creativity, and musical achievement. *Psychology of Music*, 34(1), 47–61. <https://doi.org/10.1177/0305735606059104>
- Primi, R. (2014). Divergent productions of metaphors: Combining many-facet Rasch measurement and cognitive psychology in the assessment of creativity. *Psychology of Aesthetics, Creativity, and the Arts* (online publication).
<https://doi.org/10.1037/a0038055>
- Puryear, J. S., & Lamb, K. N. (2020). Defining creativity: How far have we come since Plucker, Beghetto, and Dow? *Creativity Research Journal*, 32(3), 206–214.
<https://doi.org/10.1080/10400419.2020.1821552>
- Qian, M., Plucker, J. A., & Yang, X. (2019). Is creativity domain specific or domain general? Evidence from multilevel explanatory item response theory models. *Thinking Skills and Creativity*, 33, 100571-.
<https://doi.org/10.1016/j.tsc.2019.100571>
- Qin, J. (2024). Relationship between learner creativity, task planning, and speaking performance in L2 narrative tasks. *International Review of Applied Linguistics in Language Teaching, IRAL*. <https://doi.org/10.1515/iral-2023-0252>
- Rabab’ah, G. (2005). Communication Problems Facing Arab Learners of English. *Grazer Linguistische Studien*, 63(spring), 63–75.
- Rahi, S. (2017). Research design and methods: A systematic review of research paradigms, sampling issues and instruments development. *International Journal of Economics & Management Sciences*, 6(2), 1-5.
<https://dx.doi.org/10.4172/2162-6359.1000403>

- Reis, S. M., & Renzulli, J. S. (1991). The assessment of creative products in programs for gifted and talented students. *Gifted Child Quarterly*, 35(3), 128–134.
<https://doi.org/10.1177/001698629103500304>
- Reiter-Palmon, R., Robinson-Morrall, E. J., Kaufman, J. C., & Santo, J. B. (2012). Evaluation of self-perceptions of creativity: Is it a useful criterion? *Creativity Research Journal*, 24, 107-114. <https://doi.org/10.1080/10400419.2012.676980>
- Renandya, W. A., & Nguyen, M. T. T. (2023). Teaching speaking in L2 contexts. In *Handbook of Practical Second Language Teaching and Learning* (1st ed., Vol. 1, pp. 269–280). Routledge. <https://doi.org/10.4324/9781003106609-22>
- Renzulli, J. S. (1994). *Schools for talent development: A practical plan for total school improvement*. Mansfield Center, CT: Creative Learning Press.
- Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan*, 42, 109-135.
- Richards, J. C. (2008). *Teaching listening and speaking* (Vol. 35). Cambridge, England: Cambridge university press.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge university press.
- Ritchie, S. J., Luciano, M., Hansell, N. K., Wright, M. J., & Bates, T. C. (2013). The relationship of reading ability to creativity: Positive, not negative associations. *Learning and Individual Differences*, 26, 171-176.
<https://doi.org/10.1016/j.lindif.2013.02.009>
- Roemmele, M. (2021). Inspiration through observation: Demonstrating the influence of automatically generated text on creative writing. *arXiv preprint arXiv:2107.04007*.
- Runco, M. A., & Smith, W. R. (1992). Interpersonal and intrapersonal evaluations of creative ideas. *Personality and Individual Differences*, 13(3), 295–302.
[https://doi.org/10.1016/0191-8869\(92\)90105-X](https://doi.org/10.1016/0191-8869(92)90105-X)
- Runco, M. A., Acar, S., & Cayirdaga, N. (2017). A closer look at the creativity gap and why students are less creative at school than outside of school. *Thinking Skills and Creativity*, 24, 242–249. <https://doi.org/10.1016/j.tsc.2017.04.003>

- Said-Metwaly, S., Noortgate, W. & Kyndt, E. (2017). Approaches to Measuring Creativity: A Systematic Literature Review. *Creativity. Theories – Research - Applications*, 4(2) 238-275. <https://doi.org/10.1515/ctra-2017-0013>
- Sawyer, K. (2012). Extending Sociocultural Theory to Group Creativity. *Vocations and Learning*, 5(1), 59–75. <https://doi.org/10.1007/s12186-011-9066-5>
- 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>
- Shields, L., & Smyth, W. (2016). Common quantitative methods. In Z. Schneider, D. Whitehead, G. Lobiondo-Wood & J. Harber (Eds.), *Nursing and midwifery research: Methods, critical appraisal and utilisation* (5th ed.). Sydney, NSW: Mosby.
- Silvia, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychological Review* (online publication). <https://doi.org/10.1007/s10648-015-9299-1>
- 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., Nusbaum, E. C., Berg, C., Martin, C., & O'Connor, A. (2009). Openness to experience, plasticity, and creativity: Exploring lower-order, high-order, and interactive effects. *Journal of Research in Personality*, 43(6), 1087–1090. <https://doi.org/10.1016/j.jrp.2009.04.015>
- Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., Martinez, J. L., & 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), 68–85. <https://doi.org/10.1037/1931-3896.2.2.68>
- Simonton, D. K. (2007). The creative process in Picasso's Guernica sketches: Monotonic improvements versus nonmonotonic variants. *Creativity Research Journal*, 19(4), 329–344. <https://doi.org/10.1080/10400410701753291>

- Simonton, D. K. (2012). Taking the U.S. Patent Office criteria seriously: A quantitative three-criterion creativity definition and its implications. *Creativity Research Journal*, 24 (2–3), 97–106. <https://doi.org/10.1080/10400419.2012.676974>
- Simonton, D. K. (2016). Defining creativity: Don't we also need to define what is not creative? *The Journal of Creative Behavior*, 52(1), 80–90. <https://doi.org/10.1002/jocb.137>
- Sio, U. N., & Ormerod, T. C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, 135, 94–120. <http://dx.doi.org/10.1037/a0014212>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(1), 1–14. <https://doi.org/10.1017/S026144480200188X>
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532. <https://doi.org/10.1093/applin/amp047>
- Smith, C. A. (2013). Student creativity and language performance. In N. Sonda, & A. Krause (Eds.), *JALT 2013 conference proceedings* (pp. 285–297). Tokyo: JALT.
- Smith, G. F. (1998). Idea generation techniques: A formulary of active ingredients. *Journal of Creative Behavior*, 32, 107–134. <https://doi.org/10.1002/j.2162-6057.1998.tb00810.x>
- Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, 113(1), 117–143. <https://doi.org/10.1037/pspp0000096>
- Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within-group analysis approach. *Frontiers in Psychology*, 2, 1–8. <https://doi.org/10.3389/fpsyg.2011.00183>
- Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a

- within group analysis approach. *Frontiers in Psychology*, 2, 183.
<https://doi.org/10.3389/fpsyg.2011.00183>
- Sternberg, R. J. (2002). “Creativity as a decision”: Comment. *American Psychologist*, 57(5), 376. <https://doi.org/10.1037/0003-066X.57.5.376a>
- Sternberg, R. J. (2010). Teaching for creativity. In R. A. Beghetto & J. C. Kaufman (Eds.), *Nurturing creativity in the classroom* (pp. 394–414). Cambridge University Press. <https://doi.org/10.1017/CBO9780511781629.020>
- Sternberg, R. J. (2018). The triangle of creativity. In R. J. Sternberg & J. C. Kaufman (Eds.), *The nature of human creativity* (pp. 318–334). Cambridge University Press. <https://doi.org/10.1017/9781108185936.023>
- Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. Free Press.
- Stillar, S. (2013). Raising critical consciousness via creative writing in the EFL classroom. *TESOL Journal*, 4(1), 164–174. <https://doi.org/10.1002/tesj.67>
- Sturm, J. A. & Seery, C. H. (2007). Speech and articulatory rates of school-age children in conversation and narrative contexts. *Language and Speech and Hearing Services in Schools*, 38, 47–59. [https://doi.org/10.1044/0161-1461\(2007/005\)](https://doi.org/10.1044/0161-1461(2007/005))
- Sun, Y. C. (2009). Using a two-tier test in examining Taiwan graduate students’ perspectives on paraphrasing strategies. *Asia Pacific Education Review*, 10(3), 399–408. <https://doi.org/10.1007/s12564-009-9035-y>
- Suzuki, S., Kormos, J., & Uchihara, T. (2021). The relationship between utterance and perceived fluency: A meta-analysis of correlational studies. *The Modern Language Journal*, 105(2), 435–463. <https://doi.org/10.1111/modl.12706>
- Suzuki, S., Yasuda, T., Hanzawa, K., & Kormos, J. (2022). How Does Creativity Affect Second Language Speech Production? The Moderating Role of Speaking Task Type. *TESOL Quarterly*, 56(4), 1320–1344. <https://doi.org/10.1002/tesq.3104>
- Takeuchi, H., Taki, Y., Hashizume, H., Sassa, Y., Nagase, T., Nouchi, R., & Kawashima, R. (2011). Failing to deactivate: the association between brain activity during a working memory task and creativity. *NeuroImage*, 55(2), 681–687. <https://doi.org/10.1016/j.neuroimage.2010.11.052>

- Tan, M., Mourgues, C., Hein, S., McCormick, J., Barbot, B., & Grigorenko, E. (2015). Differences in judgments of creativity: How do academic domain, personality, and self-reported creativity influence novice judges' evaluations of creative productions? *Journal of Intelligence*, 3(3), 73–90.
<https://doi.org/10.3390/jintelligence3030073>.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–273). Amsterdam, Netherlands: John Benjamins.
- Taylor, C. W., & Barron, F. (Eds.). (1963). *Scientific creativity: Its recognition and development*. John Wiley.
- Taylor, L., & Galaczi, E. (2011). Scoring validity. *Examining speaking: Research and practice in assessing second language speaking*, 30, 171-233.
- Thornbury, S. (2005). *How to teach speaking*. Longman.
- Thornbury, S., & Slade, D. (2006). *Conversation: From description to pedagogy*. Cambridge University Press.
- Torrance, E. P. (1962). *Guiding Creative Talent*. Englewood Cliffs, NJ: Prentice-Hall, Inc. <http://dx.doi.org/10.1037/13134-000>
- Torrance, E. P. (1966). *Torrance tests of creative thinking—norms technical manual research edition—verbal tests, forms A and B—figural tests, forms A and B*. Princeton: Personnel Press. Inc.
- Torrance, E. P. (1974). *Torrance Tests of Creative Thinking: Norms-technical manual*. Bensenville, IL: Scholastic Testing Service.
- Torrance, E. P. (1982). Hemisphericity and creative functioning. *Journal of Research & Development in Education*, 15(3), 29–37.
- Torrance, E. P. (2008). *Torrance Tests of Creative Thinking: Norms-technical manual, verbal forms A and B*. Bensenville, IL: Scholastic Testing Service.
- Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge University Press.
- Ur, P. (1996). *A course in Language Teaching. Practice and Theory*. Cambridge: Cambridge University Press.

- van de Kamp, M. T., Admiraal, W., van Drie, J., & Rijlaarsdam, G. (2015). Enhancing divergent thinking in visual arts education: Effects of explicit instruction of meta-cognition. *The British journal of educational psychology*, 85(1), 47–58. <https://doi.org/10.1111/bjep.12061>
- Vartanian, O. (2009). Variable attention facilitates creative problem solving. *Psychology of Aesthetics, Creativity, and the Arts*, 3(1), 57–59. <https://doi.org/10.1037/a0014781>
- Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., & Duyck, W. (2016). The influence of language-switching experience on the bilingual executive control advantage. *Bilingualism: Language and Cognition*, 19(1), 181–190. <https://doi.org/10.1017/S1366728914000352>
- Wallach, M. A. & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity-intelligence distinction*. New York: Holt, Rinehart and Winston.
- Wang, A. Y. (2012). Exploring the relationship of creative thinking to reading and writing. *Thinking skills and creativity*, 7(1), 38-47. <https://doi.org/10.1016/j.tsc.2011.09.001>
- Wang, A. Y. (2012). Exploring the relationship of creative thinking to reading and writing. *Thinking skills and creativity*, 7(1), 38-47. <https://doi.org/10.1016/j.tsc.2011.09.001>
- Wang, H. C. (2021). Exploring the relationships of achievement motivation and state anxiety to creative writing performance in English as a foreign language. *Thinking Skills and Creativity*, 42, 1-8. <https://doi.org/10.1016/j.tsc.2021.100948>
- Wang, L. (2019). *Rethinking the Significance of Creative Writing: A Neglected Art Form behind the Language Learning Curriculum*, 6, 110-122. <https://doi.org/10.17863/CAM.46573>
- Wati, N. S. (2019). The effect of creativity toward students' achievement in writing ability. *Pedagogy: Journal of English Language Teaching*, 6(2), 141-147. <https://doi.org/10.32332/pedagogy.v6i2.1330>
- Weir, C. (1988). *Communicative Language Testing*. Exeter University Press.

- Weir, Cyril J., Ivana Vidaković & Evelina Galaczi. (2013). *Measured constructs: A history of Cambridge English language examinations 1913-2012*. Cambridge: Cambridge University Press.
- Weisberg, R. W., & Hass, R. (2007). We are all partly right: Comment on Simonton. *Creativity Research Journal*, 19(4), 345–360.
<https://doi.org/10.1080/10400410701753309>
- Weiss, S., & Wilhelm, O. (2022). Is Flexibility More than Fluency and Originality? *Journal of Intelligence*, 10(4), 96-.
<https://doi.org/10.3390/jintelligence10040096>
- Williams, F. E. (1980). *Creativity assessment packet*. Buffalo, NY: DOK Publishers.
- Wolf, K. M. & Mieg, H. A. (2010). Cognitive determinants of the success of inventors: Complex problem solving and deliberate use of divergent and convergent thinking. *European Journal of Cognitive Psychology*, 22, 443–462.
<https://doi.org/10.1080/09541440902916845>
- Wulandari, E. (2016, October 22-23). *Interaction between Students' Creativity and their Speaking Skill* [Paper presentation]. International Conference on Elementary and Teacher Education (ICETE) 2016 Lombok, Indonesia.
https://eprints.hamzanwadi.ac.id/4118/1/Baiq%20Fatmawati_U.H_%20ICTEE_2016.pdf
- Yang, Y., Wu, S., Duñabeitia, J. A., Jiang, K., & Li, Y. (2021). The influence of L2 proficiency on bilinguals' creativity: The key role of adaptive emotion regulation strategies during the COVID-19 Pandemic. *Frontiers in Psychology*, 12, 695014.
<https://doi.org/10.3389/fpsyg.2021.695014>
- Yang, Z., Ma, W., Cui, Y., Ye, J., Che, W., & Wang, S. (2021). Bilingual alignment pre-training for zero-shot cross-lingual transfer. *arXiv preprint arXiv:2106.01732*.
- Yeh, C.-C. (2017). Creative writing in an EFL writing class: Student perspectives. *English Teaching & Learning*, 41(3), 1-29.
<https://doi.org/10.6330/ETL.2017.41.3.01>

- Zabelina, D. L., & Robinson, M. D. (2010). Creativity as flexible cognitive control. *Psychology of Aesthetics, Creativity, and the Arts*, 4(3), 136–143.
<https://doi.org/10.1037/a0017379>
- Zeng, L., Proctor, R. W., & Salvendy, G. (2011). Can traditional divergent thinking tests be trusted in measuring and predicting real-world creativity? *Creativity Research Journal*, 23, 24–37. doi: 10.1080/10400419.2011.545713
- Zhang, B. (2013). An analysis of spoken language and written language and how they affect English language learning and teaching. *Journal of Language Teaching and Research*, 4(4), 834–838. <https://doi.org/10.4304/jltr.4.4.834-838>
- Zhang, S. (2009). The role of input, interaction, and output in the development of oral fluency. *English Language Teaching*, 2(4), 91–100.
<http://dx.doi.org/10.5539/elt.v2n4p91>
- Zhao, Y. (2014). L2 creative writers' identities: Their cognitive writing processes and sense of empowerment. *New Writing: International Journal for the Practice and Theory of Creative Writing*, 11(3), 452–466.
<https://doi.org/10.1080/14790726.2014.956124>
- Zhou, N. et al. (2004). *How English As A Second Language Affects Chinese Students Giving Presentations During Class In U.S. A*, [Thesis Proposal]. Graduate School of Marietta College.
- Zuhriyah, M., Agustina, R.K. & Fajarina, M. (2018). The Influence of Students' Creativity to Construct Sentences toward Their Speaking Skill. *Register Journal*, 11(1), pp 1- 18. <https://doi.org/10.18326/rgt.v11i1.1-18>

APPENDICES



Appendix 1. Ethics Committee Approval

MUĞLA SITKI KOÇMAN ÜNİVERSİTESİ SOSYAL VE BEŞERİ BİLİMLER ARAŞTIRMALARI ETİK KURULU - 2 KARARI

Protokol No : 230078

Karar No : 74

Araştırma Yürütücüsü	Yüksek Lisans Öğrencisi YUNUS EMRE KAHVECİ
Kurumu / Birimi	MUĞLA SITKI KOÇMAN ÜNİVERSİTESİ / İNGİLİZ DİLİ EĞİTİMİ
Araştırmanın Başlığı	Creativity In Foreign Language Speech Production: Investigating The Role Of Time Allocation And Task Type
Başvuru Formunun Etik Kurula Geldiği Tarih	02.08.2023
Başvuru Formunun Etik Kurulda İncelendiği Tarih	03.08.2023
Karar Tarihi	14.08.2023

KARAR : **UYGUNDUR**

AÇIKLAMA :Beyan edilen veri formlarının dışına çıkılmaması şartıyla araştırmanın uygulanabilirliği konusunda bilimsel araştırmalar etiği açısından bir sakınca yoktur.

Prof.Dr. Nilsun SARIYER
Başkan

Doç.Dr. Pınar YÜRÜR
Üye

Prof.Dr. Serkan ÇİÇEK
Üye

Doç.Dr. Perihan KORKUT
Üye
Muğla Sıtkı Koçman Üniversitesi Bilimsel Araştırmalar Etik Kurulunun Yönergesinin Madde-8/6-b bendine göre değerlendirilmediği yer almamıştır.

Doç.Dr. Raşit AVCI
Üye

Prof.Dr. Oktay YİVLİ
Üye

Doç.Dr. Savaş ARTUĞER
Üye

Prof. Dr. Tonguç Osman MUTLU
Üye

Bu belge elektronik imza ile imzalanmıştır.
Doğrulama adresi <https://etikkurulbasvuru.mu.edu.tr/dogrulama/CAJV7858>



Appendix 2. Personal Information Form**Katılımcı Bilgi Formu**

Adınız:

Soyadınız:

Yaşınız:

Mezun olduğunuz lise türü:

Dil öğrenme geçmişiniz:

(İngilizce öğrenmeye kaç yaşında, nerede, nasıl başladınız?)

Kendinizi İngilizce konusunda başarılı buluyor musunuz, neden?

İngilizce öğretmenliği bölümünü tercih etme nedeniniz?

Appendix 3. *Permission to Use RAT*

GA
Kime: Siz

UBT SETLER CVP ANAHTARL...
2 MB

Merhaba Emre,

Çok güzel ve iddialı bir konu seçmişsiniz, tez çalışmanızda şimdiden başanlar dileyerek başlayayım cevabıma. Yaratıcılığın konuşma üzerindeki etkisini inceleyeceğinizi yazmışsınız, Psikoloji Bölümlerinde araştırma yöntemleri veya istatistik içerikli dersler olur, o nedenle öğreniler bilimsel ifadeleri yerinde kullanırlar. Siz bu gibi dersleri aldınız mı bilmiyorum ama "etki" kelimesinin bilimsel alanda nasıl bir karşılığı olduğunun farkında olduğunuzu ve deney deseninizi de ona göre titizlikle tasarlayacağınızı varsayıyorum. Ayrıca seçtiğiniz her iki değişkene etki etme olasılığı yüksek olan diğer karıştııcı değişkenleri literatür taraması ile tespit edip kontrol etmeniz iyi olacaktır. Yoksa neyin neyi etkilediği; diğer bir ifadeyle ölçülen etkinin hangi etkene bağlı olarak ortaya çıktığı gösterilemezse sonrasında bulguları bilimsel açıdan tartışmak zor ve anlamsız olur.

Yaratıcılığı UBT ile ölçmek istediğinizi belirtmişsiniz, elbette kullanabildiğiniz memnun oluruz. Ekte UBT setlerini gönderiyorum. Biz katılımcılara sunduğumuz üçlü kelimelerin art arda sıralanmasının üçlü kelimelerin doğru çözülme ihtimaline etki edip etmediğini kontrol etmek için (diğer bir ifadeyle "sıra etkisini" bertaraf etmek için) Türkçe UBT'yi A-B-C-D olmak üzere 4 farklı set halinde hazırlamıştık. Bunu nasıl yaptığımızı makalemizde açıklamıştık zaten ama yine de neden 4 tane dosya gönderdiğimi belirtmek istedim.

Ayrıca Türkçe UBT'yi hazırlarken hem orijinalinden (Mesnick, 1962) hem de şu ana kadar yapılmış diğer çeviri/standardizasyon çalışmalarında yapılanlardan farklı bir kelime seçimi de yaptık. Makaleyi Türkçe yazdığımız için (psikoloji literatürü daha çok İngilizce olduğu için demeliyim) bu teste yaptığımız bu katkı maalesef alanda fark edilip tartışılmadı. Şöyle ki UBT'nin maddelerini oluşturmak adına Türkçe kelimeleri seçip üçlü hale getirirken biz çözüm kelimesinin o maddeye ait olan üç kelimenin de sonuna ya da önüne geliyor olmasına dikkat ederek maddeleri tasarladık. Buna göre 120 maddenin 60 tanesi ile ileriye doğru, diğer 60 tanesi ile de geriye doğru UBT puanı hesaplanabiliyor. 120 maddenin toplamı zaten toplam UBT puanı olarak hesaplanıyor. Maddeleri oluşturan kelimelerin dikkat ettiğimiz bu özelliği, doğru çözümü yani her bir kelime için ortak olan kelimenin doğru şekilde bulunmasını etkiliyor gibi görünüyor. Makalemizin bulguları ileriye doğru UBT maddelerinin doğru çözüm oranının geriye doğru UBT maddelerinden daha yüksek olduğunu gösteriyor. Bu ilginç bir bulgudur aslında ve sonraki çalışmalarla desteklenmesi ya da yanlışlanması anlamlı olacaktır. Siz dil alanında olduğunuz için verilerinizi maddelerin bu dil özellikleri açısından inceleyip yorumlayabilirsiniz diye özellikle belirtmek istedim.  İlim geleceği şimdilik bunlar oldu, sormak istediğiniz bir şey olursa her zaman sorabilirsiniz.

İyi çalışmalar diliyorum.



Appendix 4. Sample RAT

ÖRNEK

ELBİSE / OYUNU / RUHU



Elbise/Oyunu/Ruhu 3'lü kelimeleri için **ortak kelime 'TAKIM'** kelimesidir:

Takım elbise

Takım oyunu

Takım ruhu

Appendix 5. Sample RAT Answer Sheet

UBT Cevap Kağıdı

Tarih : 16.11.2023

Ad- soyad : [REDACTED]

Cinsiyet : Erkek

Set : b

Örnek : TAKIM

A1ıştırma a) : Gök

A1ıştırma b) : Araba

1. dışı	+	41. cı	81.		
2. ötesi	+	42.	82.		
3. son	+	43.	83.		
4.		44. Ateş	84. taş	+	
5. boş	+	45. Rüzgâr	85.		
6.		46. filmi	86.		
7. sız		47. beş	+	87. alması	
8. aşk		48. sığacı		88. arası	+
9. öz	+	49. dik	+	89. tabaka	
10. kalın		50. ağaç	+	90.	
11.		51. yeni		91. tarafı	
12. tadı	+	52. yanlış	+	92. ilk	
13. bacağı	+	53. kora	+	93. suz	
14. Ak	+	54. ön	+	94. oynar	
15. Parmak	+	55. ön	+	95. göz	+
16. dağ	+	56. yıldız		96. ada	
17. garajı	+	57. kapalı	+	97. özel	+
18. Neme	+	58.		98. nemi	
19. ilk		59. ön		99. lı	
20. yeni		60. saatı		100. italyan	
21. sız		61. geri	+	101.	
22.		62. sıra	+	102. kasaba	
23. kızıl		63. yakın	+	103.	
24. kuşu	+	64. dan		104. madde	+
25.		65. boş	+	105.	
26. son		66.		106. su	+
27. ilk		67. sever	+	107. elektrik	
28. kuvvet	+	68. zekâ	+	108. bezi	+
29.		69.		109. değer	+
30. kurumu		70. bütü	+	110. yan	+
31. milli		71. gelir	+	111. is	+
32. iğet		72. kan	+	112. ilan	+
33. geniş	+	73. etkî		113. aracı	+
34. kutu	+	74.		114. aralı.ğı	
35. Arısı	+	75. bombası	+	115. altı	+
36. Altı	+	76. başı	+	116.	
37. sabunu	+	77. evi	+	117. deli	+
38. porseli	+	78.		118. arka	+
39. yinek	+	79.		119.	
40. bacağı	+	80. taşı	+	120. tel	+

Appendix 6. BFI-2 Personality Inventory

Aşağıda sizi kısmen tanımlayan (ya da pek tanımlayamayan) bir takım özellikler sunulmaktadır. Örneğin, başkaları ile zaman geçirmekten hoşlanan birisi olduğunuzu düşünüyor musunuz? Lütfen aşağıda verilen özelliklerin sizi ne oranda yansıttığını ya da yansıtmadığını belirtmek için sizi en iyi tanımlayan ifadeyi işaretleyiniz.

Kendimi biri olarak görüyorum.

		Hiç katılmıyorum	Biraz katılmıyorum	Ne katılıyorum ne de katılmıyorum	Biraz katılıyorum	Tamamen katılıyorum
1	Dışadönük, sosyal	1	2	3	4	5
2	Şefkatli, yumuşak kalpli	1	2	3	4	5
3	Dağınık olma eğiliminde	1	2	3	4	5
4	Rahat, stresle baş edebilen	1	2	3	4	5
5	Sanatsal ilgileri az olan	1	2	3	4	5
6	Atılgan, girişken	1	2	3	4	5
7	Saygılı, başkalarına saygılı davranan	1	2	3	4	5
8	Tembelliğe eğilimli	1	2	3	4	5
9	Bir aksilik yaşadığında iyimserliğini koruyan	1	2	3	4	5
10	Farklı birçok şeye merak duyan	1	2	3	4	5
11	Nadiren heyecanlanan ya da heveslenen	1	2	3	4	5
12	Başkalarında hata arama eğiliminde olan	1	2	3	4	5
13	Güvenilir, istikrarlı	1	2	3	4	5
14	Dakikası dakikasına uymayan, ruh hali inişli çıkışlı	1	2	3	4	5
15	Yaratıcı, bir işi yapmanın akılcıca yöntemlerini bulan	1	2	3	4	5
16	Sessiz olmaya eğilimli	1	2	3	4	5
17	Başkalarının halinden pek anlamayan	1	2	3	4	5
18	Sistemli, her şeyin düzenli olmasını seven	1	2	3	4	5
19	Gergin olabilen	1	2	3	4	5
20	Sanat, müzik ya da edebiyatla çok ilgili	1	2	3	4	5
21	Baskın, lider gibi davranan	1	2	3	4	5
22	Başkaları ile tartışma başlatan	1	2	3	4	5
23	İşe başlamakta zorlanan	1	2	3	4	5
24	Güvenli, kendiyle barışık	1	2	3	4	5
25	Entelektüel, felsefi tartışmalardan kaçınan	1	2	3	4	5
26	Başkalarından daha az hareketli	1	2	3	4	5

BFI-2 Personality Inventory (continued)

27	Affedici bir yapısı olan	1	2	3	4	5
28	Biraz özensiz olabilen	1	2	3	4	5
29	Duygusal olarak dengeli, keyfi kolay kaçımayan	1	2	3	4	5
30	Yaratıcı yönü zayıf olan	1	2	3	4	5
31	Bazen utangaç, içe dönük	1	2	3	4	5
32	Yardımsever, bencil olmayan	1	2	3	4	5
33	Etrafını temiz ve derli toplu tutan	1	2	3	4	5
34	Çok endişelenen	1	2	3	4	5
35	Sanata ve estetiğe değer veren	1	2	3	4	5
36	Başkalarını etkilemede zorlanan	1	2	3	4	5
37	Zaman zaman başkalarına kaba davranan	1	2	3	4	5
38	Verimli, iş bitiren	1	2	3	4	5
39	Sıkça üzgün hisseden	1	2	3	4	5
40	Çok yönlü, derin düşünen	1	2	3	4	5
41	Enerji dolu	1	2	3	4	5
42	Başkalarının iyi niyetinden şüphe eden	1	2	3	4	5
43	Sözünde duran, başkalarının güvenebildiği	1	2	3	4	5
44	Duygularını kontrol altında tutan	1	2	3	4	5
45	Zihinde canlandırma yapmada zorlanan	1	2	3	4	5
46	Konuşkan	1	2	3	4	5
47	Soğuk ve başkalarını umursamayan	1	2	3	4	5
48	Arkasını toplamayan, dağınık bırakan	1	2	3	4	5
49	Nadiren kaygılanan ya da korkan	1	2	3	4	5
50	Şiir ve tiyatroyu sıkıca bulan	1	2	3	4	5
51	Kararları başkalarının vermesini tercih eden	1	2	3	4	5
52	Kibar, başkalarına nezaketle yaklaşan	1	2	3	4	5
53	Kolay vazgeçmeyen, işin sonunu getiren	1	2	3	4	5
54	Depresif, hüzünlü hissetmeye eğilimli	1	2	3	4	5
55	Soyut konulara az ilgi duyan	1	2	3	4	5
56	Coşku dolu	1	2	3	4	5
57	Başkaları hakkında hep iyi düşünen	1	2	3	4	5
58	Bazen sorumsuzca davranan	1	2	3	4	5
59	Değişken mizaçlı, çabuk sinirlenen	1	2	3	4	5
60	Özgün, yeni fikirler üreten	1	2	3	4	5

Appendix 7. Sample AT Prompts (*The Official Guide to the TOEFL Test*)

13. Some people believe that the best way to succeed is to set goals and work hard to achieve them. Other people think that hard work is not as important as good luck. Which point of view do you think is true and why? Use specific reasons and examples to support your opinion.
14. Some people are attracted to jobs that include a great deal of travel. Other people prefer jobs that allow them to return to their homes every evening. Which type of job opportunity would you prefer and why? Use specific reasons and examples to support your opinion.
15. Think about a book that you have enjoyed reading. Why did you like it? What was especially interesting about the book? Use specific details and examples to support your response.
16. Agree or disagree with the following statement:
Traveling independently is better than traveling as part of a tour group.
Use specific reasons and examples to support your opinion.
17. Some people who unexpectedly receive a large amount of money spend it on practical things, while others spend it for pleasure only. Which do you think is better and why?
18. Some students would prefer to live with roommates. Others would prefer to live alone. Which option would you prefer and why?
19. Some students prefer to work on class assignments by themselves. Others believe it is better to work in a group. Which do you prefer? Explain why.
20. Do you agree or disagree with the following statement?
It is important to learn about other cultures.
Use details and examples to explain your opinion.
21. Talk about a place you enjoyed going to or visiting when you were a child. Describe the place. Explain why you enjoyed it.
22. Do you agree or disagree with the following statement? Why or why not? Use details and examples to explain your answer.
It is more important to study math or science than it is to study art or literature.

Appendix 8. Sample PN Task (Adapted From Campbell, 2022)



Appendix 9. Sample Unpruned Transcription

1 ID42 Transcript

2 AT

3 I prefer to use my spare time for movies and television. All the TV series. Because there are a
4 lot more movies, good for my taste. I mean the science fiction, and I like to see things, good
5 cgi, and things about space, things I can't imagine myself very well. They do it good, Marvel,
6 for example. And I'm not really watching TV, I just watch things online on Netflix or Amazon
7 Prime. And books and magazines are a bit boring for me and I get bored easily. I'm used to
8 watching series, watching this one series for a few episodes, and watching another one for a
9 few episodes. But when I do that in books, I forget what I was reading and it is boring.
10 Maybe I just prefer reading books because they like to imagine things, or, maybe they don't
11 like to watch TV. And watching movies also improves my English. That's how I learn
12 English, actually. I started with TV series, Supernatural. And then I just learned things,
13 heard from the movies. Kept it in memory and used it.

14 PN

15 Yes, he met a beautiful girl in the bar. And asked her out for a date on Saturday. He went
16 home and his clothes were old and bad. So he went and bought some stuff, sold his car for
17 more money and cash, cut his hair down. And he wanted to impress this girl. Took her to
18 expensive restaurant. He wanted to make it look like he's rich. No problem crossing around.
19 His card was declined, so she paid with American Express, which is a big deal. That means
20 she has money. She looks a bit mad and the guy looks disappointed and sad. I think they're
21 not going on a secon- second date.



Appendix 10. Sample AS-unit Transcription

1 ID42 Transcript

2 AT

3 | I prefer to use my spare time for movies and television. | All the TV series. | Because there
 4 are a lot more movies, :: good for my taste. | I mean the science fiction, | and I like to see
 5 things, good cgi, and things about space, things I can't imagine myself very well. | They do it
 6 good, Marvel, for example. | And I'm not really watching TV, | I just watch things online on
 7 Netflix or Amazon Prime. | And books and magazines are a bit boring for me | and I get bored
 8 easily. | I'm used to watching series, watching this one series for a few episodes, and watching
 9 another one for a few episodes. | But when I do that in books, :: I forget what I was reading |
 10 and it is boring. | Maybe I just prefer reading books :: because they like to imagine things, | or,
 11 maybe they don't like to watch TV. | And watching movies also improves my English. | That's
 12 how I learn English, actually. | I started with TV series, Supernatural. | And then I just learned
 13 things, :: heard from the movies. | Kept it in memory and used it. |

14 PN

15 | Yes, he met a beautiful girl in the bar. | And asked her out for a date on Saturday. | He went
 16 home and his clothes were old and bad. | So he went and bought some stuff, | sold his car for
 17 more money and cash, | cut his hair down. | And he wanted to impress this girl. | Took her to
 18 expensive restaurant. | He wanted to make it look like he's rich. | No problem crossing around.
 19 | His card was declined, so she paid with American Express, :: which is a big deal. | That
 20 means she has money. | She looks a bit mad | and the guy looks disappointed and sad. | I think
 21 they're not going on a {secon-} second date. |



Appendix 11. Sample Pruned Transcription

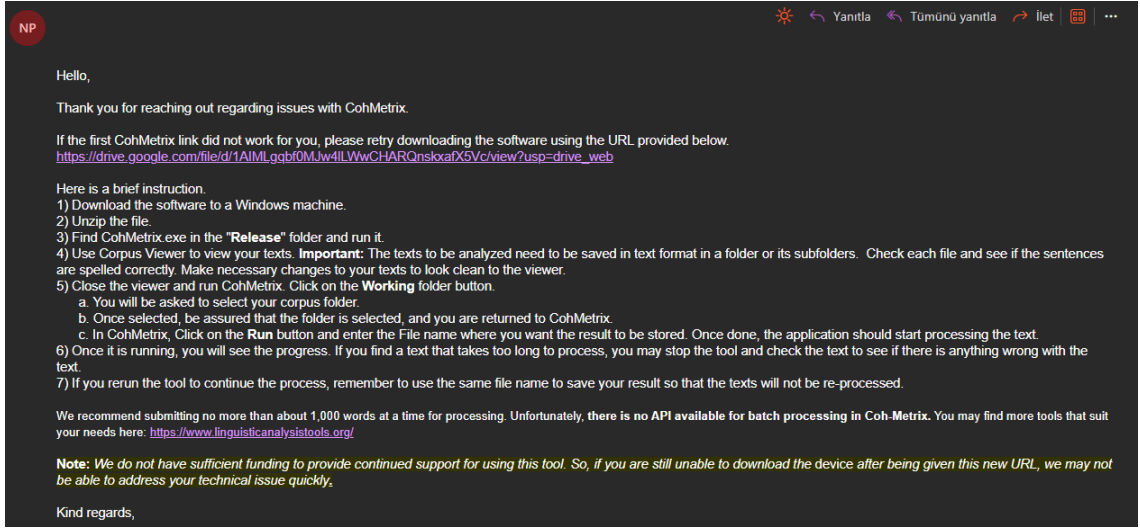
ID42

I prefer to use my spare time for movies and television. All the TV series. Because there are a lot more movies, good for my taste. I mean the science fiction, and I like to see things, good CGI, and things about space, things I can't imagine myself very well. They do it good, Marvel, for example. And I'm not really watching TV, I just watch things online on Netflix or Amazon Prime. And books and magazines are a bit boring for me and I get bored easily. I'm used to watching series, watching this one series for a few episodes, and watching another one for a few episodes. But when I do that in books, I forget what I was reading and it is boring. Maybe I just prefer reading books because they like to imagine things, or, maybe they don't like to watch TV. And watching movies also improves my English. That's how I learn English, actually. I started with TV series, Supernatural. And then I just learned things, heard from the movies. Kept it in memory and used it.

He met a beautiful girl in the bar. And asked her out for a date on Saturday. He went home and his clothes were old and bad. So he went and bought some stuff, sold his car for more money and cash, cut his hair down. And he wanted to impress this girl. Took her to expensive restaurant. He wanted to make it look like he's rich. No problem crossing around. His card was declined, so she paid with American Express, which is a big deal. That means she has money. She looks a bit mad and the guy looks disappointed and sad. I think they're not going on a second date.



Appendix 12. Permission to Use Coh-Metrix Software



Appendix 13. Sample Weighted-clause Ratio

ID18 Argumentative Task		
1. I seen the picture once time.	3	0,1
2. It had is Socrates' death.	3	0,1
3. And this is so amused me,	3	0,1
4. because this is so exciting me,	1	0,8
5. because Socrates' finger is on the head.	1	0,8
6. And there is a lot of people other him.	1	0,8
7. Because Socrates will dead after this time.	2	0,5
8. This justify is detour the Athena's children, young.	3	0,1
9. So they want to punish him.	Accurate	1
10. But Socrates was not endured against this situation.	2	0,5
11. He look like calm and sane.	1	0,8
12. Because he said that:	Accurate	1
13. "I can nothing.	3	0,1
14. I am nothing to do.	3	0,1
15. Because I am not defence false things.	2	0,5
16. It's just the government decision."	Accurate	1
17. So I like the Socrates.	1	0,8
18. I like the Socrates demean.	1	0,8
	$9,9 / 18 = .55$	9,9

Appendix 14. *Permission to Use Praat Software*

