

EFFECTS OF PROFICIENCY IN A SECOND LANGUAGE ON FALSE
MEMORIES PRODUCED BY ASSOCIATIVE AND CATEGORICAL
PROCEDURES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BAHAR AYKUT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
THE DEPARTMENT OF PSYCHOLOGY

AUGUST 2024

Approval of the thesis:

**EFFECTS OF PROFICIENCY IN A SECOND LANGUAGE ON FALSE
MEMORIES PRODUCED BY ASSOCIATIVE AND CATEGORICAL
PROCEDURES**

submitted by **BAHAR AYKUT** in partial fulfillment of the requirements for the degree
of **Master of Science in Psychology, the Graduate School of Social Sciences of Middle
East Technical University** by,

Prof. Dr. Sadettin KİRAZCI
Dean
Graduate School of Social Sciences

Prof. Dr. Bengi ÖNER ÖZKAN
Head of Department
Department of Psychology

Prof. Dr. Mine MISIRLISOY BIYIKOĞLU
Supervisor
Department of Psychology

Examining Committee Members:

Assoc. Prof. Dr. Banu CİNGÖZ ULU (Head of the Examining Committee)
Middle East Technical University
Department of Psychology

Prof. Dr. Mine MISIRLISOY BIYIKOĞLU (Supervisor)
Middle East Technical University
Department of Psychology

Assist. Prof. Dr. Özlem ERTAN KAYA
Ankara Medipol University
Department of Psychology

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Bahar AYKUT

Signature:

ABSTRACT

EFFECTS OF PROFICIENCY IN A SECOND LANGUAGE ON FALSE MEMORIES
PRODUCED BY ASSOCIATIVE AND CATEGORICAL PROCEDURES

AYKUT, Bahar

M.S., The Department of Psychology

Supervisor: Prof. Dr. Mine MISIRLISOY BIYIKOĞLU

August 2024, 64 pages

False memories have been studied extensively to understand the underlying mechanisms of various cognitive processes (Gallo, 2010). A common method to explore false memory is the Deese, Roediger, and McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). Research focusing on the effect of language on DRM-elicited false memories has shown not only that false recognition rates are higher in the dominant language (L1) compared to the non-dominant language (L2) (Anastasi et al., 2005), but also that false memory rates increase with proficiency (Arndt & Beato, 2017). These differences were explained with differential activation in L2 and L1 (Dewhurst et al., 2009). In addition, studies suggest that DRM and category lists create different activation levels (Coane et al., 2021). However, the differential effects of list type on false memories at different levels of proficiency in L2 have not yet been investigated. The current study aims to investigate the effect of proficiency in L2 on the differential false memory rates elicited by DRM and category lists. To this end, we ran a two-session experiment in which we contrasted the false memory rates elicited with DRM and category lists for individuals with differential levels of L2 proficiency. L2 proficiency was expected to affect DRM and category lists differently. Results showed that DRM and category lists differed regarding false

memory rates in L2, supporting previous literature. However, levels of L2 proficiency did not affect false memory rates in either list type, the reasons for which are discussed regarding underlying mechanisms.

Keywords: False memory, second language proficiency, DRM paradigm, category lists, recognition memory

ÖZ

İKİNCİ DİLDE DİL YETERLİLİĞİ SEVİYELERİNİN DEESE-ROEDIGER-MCDERMOTT (DRM) LİSTELERİ VE KATEGORİK ÇAĞRIŞIMLI LİSTELER İLE OLUŞTURULAN BELLEK YANILMALARINA ETKİSİ

AYKUT, Bahar

Yüksek Lisans, Psikoloji Bölümü

Danışman: Prof. Dr. Mine MISIRLISOY BIYIKOĞLU

Ağustos 2024, 64 sayfa

Yanlış bellek, çeşitli bilişsel süreçlerin temel mekanizmalarını anlamak amacıyla literatürde kapsamlı bir şekilde incelenmiştir (Gallo, 2010). Yanlış belleği araştırmak için yaygın olarak kullanılan bir yöntem Deese, Roediger ve McDermott (DRM) paradigmasıdır (Deese, 1959; Roediger & McDermott, 1995). Anadil ve ikinci dilin DRM paradigması ile üretilmiş yanlış bellek üzerindeki etkisini inceleyen araştırmalar, yanlış tanıma oranlarının baskın dilde (L1) ikinci dilden (L2) daha yüksek olduğunu ve bu etkinin dildeki yeterlilik ile arttığını göstermiştir (Anastasi ve diğ., 2005; Arndt & Beato, 2017). Bu farklılıklar, L2 ve L1'deki farklı aktivasyon seviyeleriyle açıklanmıştır (Dewhurst ve diğ., 2009). Ayrıca, çalışmalar, kategori ve DRM listelerinin farklı aktivasyon seviyeleri ortaya çıkardığını ortaya koymuştur (Coane ve diğ., 2021). Ancak, bildiğimiz kadarıyla, liste türünün ikinci dildeki farklı yeterlilik seviyelerinde yanlış bellek üzerindeki farklı etkilerini araştıran bir çalışma bulunmamaktadır. Mevcut çalışma, ikinci dildeki yeterliliğin DRM ve kategori listeleri tarafından ortaya çıkarılan farklı yanlış bellek oranları üzerindeki etkisini araştırmayı amaçlamaktadır. Bu amaçla, farklı yeterlilik seviyelerine sahip bireylerin DRM ve kategori listeleri ile elde edilen yanlış bellek oranlarının karşılaştırıldığı iki

oturumluk bir deney gerekleřtirilmiřtir. İkinci dildeki yeterlilięin DRM ve kategori listelerini farklı řekilde etkilemesi beklenmiřtir. Sonular, DRM ve kategori listelerinin, literatürdeki bulguları destekleyen řekilde, L2'deki yanlıř bellek oranları aısından farklılık gösterdięini göstermiřtir. Ancak, ikinci dildeki yeterlilik seviyeleri her iki liste türündeki yanlıř bellek oranlarını etkilememiřtir; bunun nedenleri altta yatan temel mekanizmalar aısından tartıřılmaktadır.

Anahtar Kelimeler: Yanlıř bellek, ikinci dil yeterlilięi, DRM paradigması, kategori listeleri, tanıma belleęi

To every member of my lovely family

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Prof. Dr. Mine Mısırlısoy Bıyıkođlu, for her unwavering guidance, support, and understanding throughout my journey. Her encouragement helped me navigate the challenges of returning to academia after a long hiatus. I am also sincerely appreciative of my thesis committee members, Assoc. Prof. Dr. Banu Cingöz Ulu and Assoc. Prof. Dr. Özlem Ertan Kaya, for dedicating their time to be part of my jury and for their invaluable insights.

Many thanks to Elif Kaşgöz for always being available to answer my questions about our courses and for her advice on the data collection process. I am deeply grateful to Ceyda Koca for her immense help in recruiting participants, and Abd for being an inspiration and a helping hand. I would also like to thank Aslıhan Berra Bolat, Cemile Ceren Akgöl, and Karya İlkim Ekiz for their friendship and for taking time out of their day to participate in the pilot studies.

Lastly, I extend my deepest thanks to my extended family for their unwavering support. To my parents, your unconditional love and belief in me provided the foundation for my courage to take each step forward. To my sister, you have been a constant source of joy. To my lovely grandmother, your home has always been a refuge, if it was not for you, I would not be able to chase my dreams. To my uncle, your presence during critical times has been invaluable—I would be lost without you in the city I returned to. And finally, to my dear aunt, my gamemaker, thank you for always keeping me on track, encouraging me, and lighting up even the darkest nights when sleep seemed unattainable.

TABLE OF CONTENTS

PLAGIARISM	iii
ABSTRACT	iv
ÖZ	vi
DEDICATION	viii
ACKNOWLEDGEMENTS	ix
TABLE OF CONTENTS	x
LIST OF TABLES	xii
CHAPTERS	
1.INTRODUCTION.....	1
1.1 The DRM Paradigm and the Category Associates Procedure.....	2
1.1.1 The DRM Paradigm	2
1.1.2 The Category Associates Procedure.....	3
1.1.3 Comparison of DRM and Category Associates Procedures.....	4
1.2 Accounts of False Memory	6
1.2.1 The Activation Monitoring Framework	7
1.2.2 Fuzzy Trace Theory (FTT).....	8
1.3 False Memories in a Second Language and The Role of Proficiency.....	8
2.THE CURRENT STUDY	12
2.1 Method 13	
2.1.1 Participants	13
2.1.2 Apparatus and Stimuli	14
2.1.3 Design and Procedure.....	16
3.RESULTS	19
3.1 List Type on Veridical and False Memory.....	19
3.2 Proficiency, List Type, and False Memory	21
3.2.1 Proficiency Tests	21
3.2.2 L2 Proficiency and List Type Interaction on False Memory.....	24
3.3 Proficiency and List Type on Sensivity Index (d')	26
4.DISCUSSION	30
4.1 Limitations and Future Directions.....	32
4.2 Conclusion.....	33

REFERENCES	34
APPENDICES	40
A.CATEGORY WORD LISTS	40
B.DRM LISTS	43
C.FILLER WORDS	45
D.APPROVAL OF THE METU HUMAN SUBJECTS ETHICS	46
E.INFORMED CONSENT FORM.....	47
F.TURKISH SUMMARY / TÜRKÇE ÖZET	48
G. THESIS PERMISSION FORM / TEZ İZİN FORMU.....	64

LIST OF TABLES

Table 1 Means of Yes Response Rate Based on List and Item Type	20
Table 2. Pairwise Comparisons of List and Item Type Levels.....	20
Table 4. Correlations Table For The 3 Proficiency Tests	23
Table 5. Reliability Analysis For The 3 Proficiency Tests and Their Combined Scores	23
Table 6. Item Reliability Statistics For The 3 Proficiency Tests.....	23
Table 7. Mean and SE Values For Coefficients in The Model.....	25
Table 8. Regression Coefficients For Predictors and Their Interaction on FM Rates.....	26
Table 9. Model Summary and ANOVA Tables For d' analysis	27
Table 10. Coefficients For The Covariates and The Model	27
Table 11. Model Summary and ANOVA Table For c Analysis	28

LIST OF FIGURES

Figure 1.	Representation of the Semantic Activation in Mind According to AMF	8
Figure 2.	Application of FTT to Different Events	9
Figure 3.	Illustration of the stimuli presentation during the study and the test phase.....	18
Figure 4.	The schematic illustration of the trial sequence in the study phase.....	18
Figure 5.	Mean response rates for the DRM condition across item types	21
Figure 6.	Mean response rates for the category condition across item types.....	21
Figure 7.	d' levels as a recognition sensitivity measure based on list type.	29

CHAPTER 1

INTRODUCTION

In our daily lives, we mostly think of our memories as the remnant details of the events that have taken place. However, the complex nature of systems that govern our memory can also lead to the retrieval of information for events that never took place or remembering details of these events differently; a phenomenon known as false memory (Roediger & McDermott, 1995). False memory has been studied extensively in the literature, with some studies emphasizing its forensic and clinical aspects, while others approach it as an error or illusion of memory that provides valuable information about the governing processes of both episodic and semantic memory. The present thesis will adopt the approach to false memory as a memory illusion that aids us in understanding memory processes.

Early experimental research on memory has mostly focused on the accuracy of responses. Errors were considered to be only helpful for informing about the possible criterion shifts or correcting accuracy while calculating performance (Roediger, 1996). Despite distortions and illusions of memory in and of itself not being of primal interest for most of the earlier memory researchers, there were also some studies that remarked on how those could be useful. For instance, as early as 1894, Kirkpatrick (as cited in Oliveira et al., 2018) reported that when participants were presented with words such as roll, thimble, and knife, they tended to think that related words such as thread, needle, and fork were also in the list they were presented with. He suggested that this is an important illustration of how suggestions could manipulate people's reports of an experience. Later studies employed various methods to observe memory distortions, such as the misinformation paradigm (Loftus et al., 1978), imagination inflation (Loftus et al., 1996), and the list learning paradigm (Deese, 1959). The list learning paradigm is one of the most common methods to elicit false memories in a laboratory setting. It consists of participants studying a list of items to remember for a later memory test. Studies, such as the one by Deese in 1959 and Underwood in 1965, used

lists of related words to show that when people are presented with related items, they are very likely to falsely remember another semantically related item that they were not actually presented with.

Investigating memory illusions in the lab is important to control or manipulate the various aspects that could affect the elicitation of false memories at different phases of the process. For instance, manipulations of the type of stimuli (e.g. pictorial vs. verbal), encoding and retrieval manipulations, such as imagery encoding, and priming. Observing the outcomes of such diverse manipulations is crucial for gaining a deeper understanding of the stages involved in memory processes. False memories, similar to other memory illusions, have implications about the constructive nature of memory. They can inform us on how our existing knowledge can shape the way we organize the perceived information, which hints at the effect top-down processes have on the perception of our environment as well as how we process, store, and retrieve that perceived information.

1.1 The DRM Paradigm and the Category Associates Procedure

1.1.1 The DRM Paradigm

A commonly used method of the list-learning paradigms adopted to explore false memory is the converging associates procedure developed by Deese, Roediger, and McDermott (Deese, 1959; Roediger & McDermott, 1995). In the Deese-Roediger-McDermott (DRM) paradigm, a false memory is elicited using lists of semantically related words (Gallo & Roediger, 2001). The DRM lists are associative in nature, and the words on a DRM list are not only semantically related to one another but also associated with one critical “lure” word that is not included in the study list presented to the participant. Typically, the participant mistakenly retrieves this lure word as a studied word. The retrieval rates of the non-studied critical lure (CL) words can surprisingly be very close to the correct recall or recognition rates for the actual studied items. For instance, when participants are presented with words such as pie, red, juice, and pear, all of which are related to a specific critical word that was not presented in the study, they are still very likely to retrieve this critical word, in this case apple, with the presented items. In their original study, Roediger and McDermott (1995) used Russel and Jenkins (1954) norms to form the DRM lists. Every list consisted of 12

words picked from the words given as responses to the critical lure word in associative word norms. Their results revealed that false recall rates for the critical lure words were similar to correct recall rates for studied words. For recognition memory, false alarm rates were very close to the hit rates, indicating that DRM lists were a prominent method of eliciting and investigating memory illusions in laboratory settings.

The DRM paradigm is considered to be very reliable because the false memories elicited by the DRM procedure are robust across a great number of studies that employed different encoding and retrieval strategies (Dewhurst et al, 2017), age groups (Carneiro et al., 2007), study-test intervals (Thapar & McDermott, 2001), and stimuli modality, i.e. lists using auditory, visual or pictorial stimuli (Wang et al, 2018). False memories elicited by DRM lists are often hard to distinguish from true memories for participants; they report remembering critical lure words with confidence in studies that collect confidence ratings or remember/know responses (Dewhurst & Warrand, 2004). Similarly, participants report retrieving false episodic details about the false memories, such as attributing the critical lure to a specific experimental source (Payne et al., 1996).

Despite the evidence suggesting that false memories elicited by the DRM procedure can be experienced as vividly as actual memories, some argue that the DRM false memories differ from the naturalistic creation of false memories. Therefore, generalizing the implications of these studies with DRM false memories to real-life false memories may not be appropriate (Gallo, 2010). However, research has demonstrated that individuals more susceptible to DRM-induced illusions are also more likely to experience other memory distortions or extraordinary autobiographical memories, such as those involving alien abductions (Clancy et al., 2002).

1.1.2 The Category Associates Procedure

A similar procedure used in list-learning paradigm is the Category Associates Procedure (Buchanan et al.,1999). In this procedure, the word lists are exemplars of a higher taxonomic category, and the most typical exemplar of that subordinate category happens to be the critical lure word, which is not included in the list (Dewhurst, 2001; Seamon et al., 2000). For instance, the study list in the category of fruits includes words like pear, orange, banana, etc., and the most typical member of the fruit category,

apple, is excluded from the list as the critical lure word, yet, it is still retrieved as a studied word during the test phase, making it a false memory. The false memory rates for category lists are typically lower than the false memory rates elicited by the DRM lists for both recall and recognition tests (Pierce et al., 2005; Desoto & Roediger, 2014).

Even though DRM and category lists both consist of semantically related items and they both rely on associations between the list items and the critical lure word to elicit the memory illusion, they have differential natures (Smith et al., 2000). The relationship between the list words and the critical word is only taxonomic for the category lists, indicating category membership. In contrast, for DRM lists, these relationships are multifaceted, i.e., there are multiple levels of association, such as synonymy, cooccurrence in similar contexts, shared features, etc. (Coane et al., 2021). To demonstrate this difference with the examples provided earlier, a category list with the critical word apple would include words such as pear, orange, and banana, which all belong to the same semantic category of fruits, while a DRM list with the same lure word, apple, would include words that are associatively related to apple in more than one way, such as pie (co-occurrence), red (feature-based), as well as pear (taxonomic).

1.1.3 Comparison of DRM and Category Associates Procedures

Exploring the similarities and differences between DRM and category lists helps us shed light on how false memories are generated by improving theoretical insights. When considering the similarities between DRM and category lists, a key point is that both methods utilize semantic associations to elicit false memories of words that were not actually presented. This common approach underscores how both techniques exploit the tendency of memory systems to link related concepts, resulting in activating the non-presented lure words. Studies also demonstrated that illusions elicited by DRM and category lists possibly had similarities in their susceptibility to be affected by a number of experimental manipulations. It has been argued that both for DRM and category lists, interventions that occurred during the encoding phase affected the rate of false memory illusions. For example, dividing attention during the encoding phase (Knott & Dewhurst, 2007), and explicit warning instructions given during the study about the illusion (Brainerd et al., 2001) decreased the false memory rates for both list

types. On the other hand, using blocked lists compared to randomized lists during the study phase (Tussing & Greene, 1997), using a higher number of words per list (Dewhurst & Anderson, 1999), and instructing the participants to form associations between the study items (Dewhurst et al., 2009) increased false memory rates for both DRM and category lists.

Despite their similarities, category lists and DRM lists also have key differences. While both methods give rise to false memories, they do so through different mechanisms (Coane et al., 2021). Category lists often create false memories by leveraging semantic relationships within a broader category, leading to errors based on the general meaning of the items. In contrast, DRM lists use the strength of associative links between specific words to generate false memories, focusing on individual word pairs and their connections. It has been argued that the observed differences between the two processes could stem from the differences they have in their associative nature, which further informs us about the differentiative processes that might be involved in each procedure (Knott et. al. 2012). In their 2005 study, Park, Shoebe, and Kihlstorm argued that DRM and category lists could be utilizing different processes while creating false memories because of the difference in the way the words in the lists relate to each other and to the critical word. In general, DRM lists elicit a higher proportion of false memory than category lists (Buchanan et al., 1999, Dewhurst et al., 2009). One possible reason for this finding is argued to be that the words in DRM lists have higher Backwards Associative Strength (BAS) compared to category lists (Roediger et al., 2001). Backward associative strength refers to the associative strength from the items in the study lists to the non-studied critical word (Beato & Arndt, 2017). Another study by Knott et. al. (2012) had findings that highlighted the effect of BAS on creating false memories. In both recall and recognition tasks, the proportion of memory illusions for DRM and category lists did not differ for high BAS lists and was higher than for low BAS lists. However, for low BAS lists, category lists showed no significant priming effect in word stem completion task, unlike DRM lists. On the same note, a study by Roediger and colleagues (2001) showed that a DRM list with a mean BAS of .23 (the list with king as the critical lure word) can result in a false-recall rate of .10, whereas another list with a mean BAS of .17 (the list with the word smoke as the lure), can produce a higher false recall rate (.54), despite its lower BAS level. This indicates that although BAS is a well agreed

upon measure that predicts false memory, even when it is controlled for, there are still observed differences between the two types of lists. Therefore, it is possible to infer that BAS cannot account for all the differences between the list types on its own, and there must be other factors at play as well (Coane et.al., 2016). Another difference between the DRM and category lists that have been revealed in the literature is that, when indirect priming effects were examined for DRM and category lists, priming was only observed for the DRM lists (Smith et al., 2002). The DRM procedure produces memory illusions that are prone to be influenced by associations activated at study (Roediger et al., 2001) and are not as affected by processes at test (Coane & McBride, 2006). Memory illusions produced by the category procedure are more likely to be made by semantic processes that occur at the test phase (Smith et al., 2002). Since the two processes can be affected differently by various phenomena, as well as having differences in the encoding retrieval phases and the type and rate of false memories they elicit, it was argued that these differences imply distinct underlying mechanisms of the list types (Smith et al., 2002). Therefore, to better understand how the two processes differ in terms of what they are affected by and to shed light on whether the two types of lists differ in terms of the mechanisms they employ, it could be beneficial to consider other aspects that could highlight potential differences.

1.2 Accounts of False Memory

Early research on memory illusions and disruptions, particularly involving complex narratives, explored how the mind compensates for missing information. For instance, Bartlett's (1932) study demonstrated how participants, after reading a story, recalled it with details not originally included. This phenomenon, explained by the constructive nature of memory, suggests that memory is not a mere reproduction of events but involves reconstructive processes (Brainerd & Reyna, 2005). According to constructivist theories, individuals tend to remember the meaning of an event rather than the exact details (Reyna, 1996). Bransford and Franks (1971) further proposed that shortly after an event, people lose specific details and retain an integrated representation based on their interpretations. Although false memories of real-life events and studies that utilize complex narratives can hint at the retrieval of a general interpretation rather than actual details of an event, it is also considerable that some

actual details can be stored and retrieved alongside that general meaning. Moreover, details such as the time and modality could even enhance the memory of the event. On the other hand, schema-based approaches focus on how memory is organized into schemas or networks of related concepts. This approach suggested that when individuals are exposed to a list of related words, a schema associated with that category becomes activated, spreading to related concepts, including the critical lure. This activation makes the non-presented word seem familiar and contributes to its false recognition or recall, and the fit of the critical lure within the schema contributes to the formation of false memories. Based on these earlier explanations, other theoretical frameworks attempted to explain findings in the memory illusions literature, two accounts stand out as the most frequently cited and offer explanations for a broader range of results.

1.2.1 The Activation Monitoring Framework

One of the most common attempts to explain how false memories come to be is the Activation/Monitoring Framework, or AMF, (Roediger et al., 2001), which combines activation from the Implicit Associative Response (IAR) theory of Underwood (1965), and monitoring account from the Source Monitoring Framework by Johnson (1988). According to AMF, false memories occur because the semantically related words studied in the list become activated during the study phase. This activation spreads to the non-presented critical lure word. Figure 1 represents this activation of a word and the spread of this activation to other related words in the semantic network, and to the non-presented lure word. As the amount of this activation increases, due to various factors including Backwards Associative Strength (BAS), number of words, and number of CLs, it becomes easier to accept this lure word as “old”, as in, has been presented in the study list. On the other hand, since increasing association hinders monitoring the source of this activation, it becomes harder to identify the item as “new”. As exposure and proficiency increase, so does the automaticity of activation of concepts and their associative links, making false memory production occur without conscious effort or awareness (Howe, 2005).

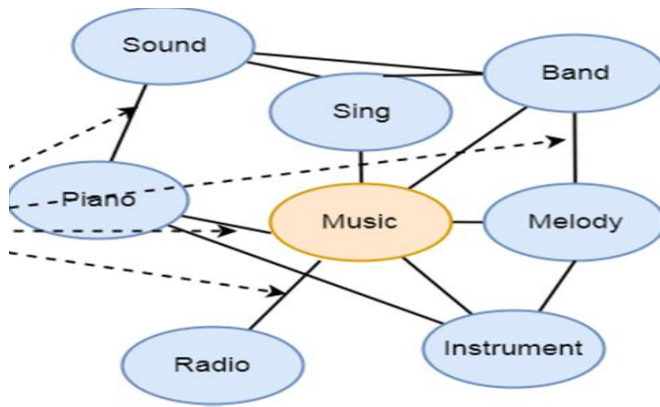


Figure 1. Representation of the Semantic Activation in Mind According to AMF

Note. The activation of the network of relevant words by the spreading activation from the word “music”. From “A self-reference false memory effect in the DRM paradigm: Evidence from Eastern and Western samples” by Wang et al, 2018, *Memory & Cognition*, 47(1), p. 77.

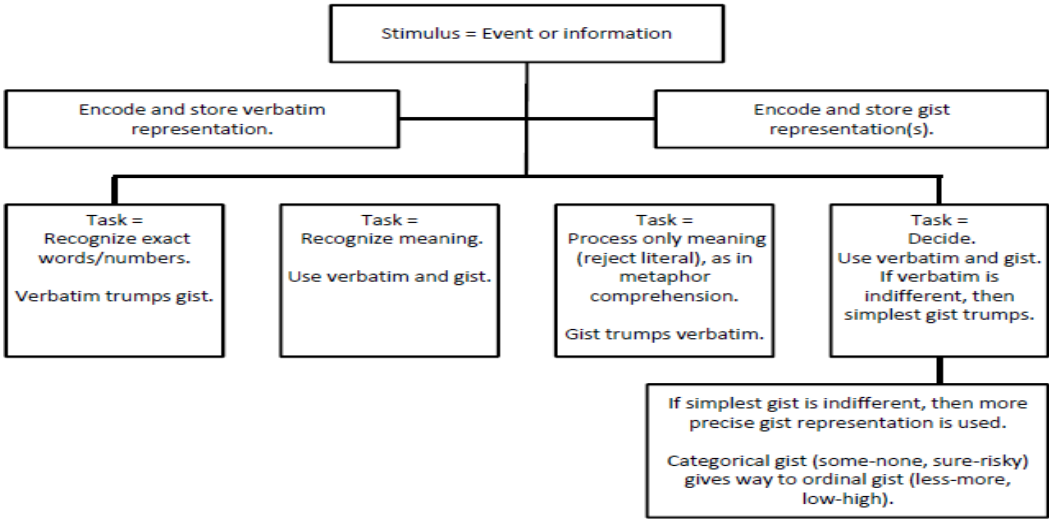
1.2.2 Fuzzy Trace Theory (FTT)

Another theoretical framework that attempts to explain false memories is Fuzzy Trace Theory (FTT), which suggests that there are 2 kinds of memory traces stored in the memory: verbatim and gist traces. Verbatim traces refer to the exact details of an event, such as the physical features of it. Gist traces, on the other hand, have the main concept of the event, such as the meaning or the main pattern (Brainerd & Reyna, 2002). For instance, when words in a list are presented, the visual aspects like the color and font of the word or the time of presentation would be verbatim details, and the general theme of the list would be the gist of it. FTT suggests that as the semantically related words in the list are studied, participants produce a gist, a summary of the general theme or meaning of the list and rely on this gist while making new/old decisions about the test words. Critical lure being closer to that created gist makes the participant more vulnerable to false memories, and retrieval of verbatim traces limits that probability. Figure 2 demonstrates how verbatim and gist traces can predominate one another in different situations.

1.3 False Memories in a Second Language and The Role of Proficiency

False memories elicited by associative lists, especially DRM lists, have been investigated in various cultures and languages; and it has been shown that DRM lists

elicit false memories in other languages as well (Misirlisoy, 2004, Turkish; Anastasi et al., 2005 Spanish; Nam, 2018, Korean).



Note. Verbatim and Gist traces predominating one another based on the characteristics of the situation. From “A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory” by Reyna, V. F., 2012, Judgment and Decision Making, 7(3), p. 337.

Figure 2. Application of FTT to Different Events

However, it is hard to compare the results of the studies investigating false memories in languages other than English for several reasons. In many studies, bilingualism and second language acquisition are used interchangeably, and their definitions are unclear. In some studies, the participants' dominant language differs from their native language (Marmolejo et al., 2009). In addition, because it is not always possible to find equivalent words in another language for the original list items, factors that could affect the level of false memories, such as BAS levels and list length (Coane et al., 2001), vary a lot between the lists used. Despite these differences between the studies, there are still some common results. In general, false memories seem to be more prominent in the first or dominant language (L1), than second or non-dominant one (L2) (Beato et al., 2023; Kawasaki-Miyaji et al., 2003; Sunderman & Kroll, 2006), and increase with proficiency (Graves et al., 2014). These common findings are accounted for by both AMF and Reverse Hierarchical Model (RHM). According to RHM, bilinguals with low proficiency access the meaning of the words through their L1 because the link between the words and their L2 is still weak; as they become more proficient in their L2, this connection gets stronger, and the mediation of L1 is less needed (Kroll & Stewart, 1994). Therefore, the weaker connection between the words

in L2 leads to lower levels of activation for the lure word. As the connections become stronger with proficiency, associations also increase, leading to a higher magnitude of false memories (Suarez & Beato, 2023). The findings in the literature regarding false memories in a second language suggest a difference in how L1 and L2 are processed in the mind. Yet, there is not enough evidence that has implications on where these differences stem from (Suarez & Beato, 2023). As mentioned earlier, existing research has divergent findings and conclusions, not allowing the derivation of implications for the mechanisms of monolingual and bilinguals' language and memory processing. For instance, because DRM lists are associatively related and the activation it creates is affected by many factors, such as BAS or frequency of the words in that specific language, using translated lists may not allow researchers to control for possible confounds (Anastasi et al., 2005; Marmolejo et al., 2009).

Researchers developed original lists in accordance with the characteristics of their own language, like frequency of use, concreteness, and imageability. However, it was not possible to control for or equalize them in every aspect as the lists in other languages (Yüvrük et al., 2018), making it difficult to compare results across languages. In addition, studies that employ only the second language for investigating false memory are also scarce. Most of the literature consists of studies that test L1 and L2 together, which prevents us from knowing clearly whether the observed effect is unique to L2 because of the possibility of unwanted activation evoked by seeing words in L1 (Suarez & Beato, 2023). Associations being weaker for concepts in L2 also places an emphasis on the effect of language proficiency on false memory, which could imply that a difference in the type and level of association could produce different results in a second language than observed so far in L1 (Arndt & Beato, 2017). Also, participants tend to be affected more by manipulations in phonological or orthographical features of the lists in their L2 compared to L1 and compared to semantic manipulations (Bialystok et al., 2020). This observation, that the participants are more affected by the phonological and orthographical feature changes in their L2, is more prominent for lower proficiency (Dewhurst & Robinson, 2004), pointing to how proficiency affects mental representations of words and semantic processes. Given that DRM and category lists use different types of associations, and create different levels of activation, and proficiency is claimed to affect activation levels as well, it's possible that language proficiency affects the false memories elicited by these methods

differently. Investigating both DRM and category lists at different L2 proficiency levels could provide important insights as to whether the typical difference between the two list types will persist in a second language, and how it would change with proficiency. This approach would help us understand how the differential levels of activation due to different levels of language proficiency impact false memory rates across different types of associative procedures and provide implications for the theoretical accounts that attempt to explain false memories. Additionally, it offers meaningful implications for understanding the distinct mechanisms underlying DRM and category lists, contributing to the ongoing debate of whether false memories elicited by DRM and category lists use the same mechanisms.

In conclusion, existing literature suggests that language proficiency affects false memory rates by enhancing the automaticity of lexical representations, which in turn activates conceptual representations in semantic memory (Howe et al., 2009). Given that DRM and category lists produce different levels of associations, examining the interaction of these factors could illuminate key theoretical debates about the nature and creation of false memory, and mental representation in a second language. This investigation not only addresses gaps in our understanding of false memories but also paves the way for future research on how proficiency in a second language influences memory. The results of this investigation between DRM and category lists for different levels of proficiency will have implications for both second language processing and memory, how they affect general differences between DRM and category lists, and the different mechanisms that govern these processes. The significance of the study is that it is the first study that compares the differential effect of proficiency in L2 for DRM and category lists.

CHAPTER 2

THE CURRENT STUDY

As mentioned in the previous chapter, DRM and category lists differ in nature, produce different rates of false memories, and are influenced differently by manipulations at encoding and retrieval and other phenomena. These differences are usually attributed to the difference in the nature of the relationship between the words each list type employs and the differential activation they create as a result of this difference in their nature.

The literature has also shown that higher proficiency in a second language results in higher proportions of false memory (Arndt & Beato, 2017), and participants tend to be affected more by manipulations in phonological or orthographical features of the lists in L2 compared to their L1. This observation is often explained by the strength of the associations between the word representations in one's L1 being stronger compared to L2, and increasing with proficiency (Dewhurst & Robinson, 2004), pointing to a difference in mental representations of words based on proficiency levels. This emphasis on association between the concepts in L2 implies that a difference in the type and level of association could produce different results for false memories in L2 than observed in L1 (Arndt & Beato, 2017). Considering the different types of associations category lists and DRM lists have, and how L1 and L2 can differ in terms of mental representation and the association between concepts, it could be possible to argue that the effect language proficiency has on the results of these two procedures could be different.

Another explanation for DRM lists producing higher levels of false memory comes from the monitoring approach, suggesting that category lists are affected more by monitoring processes than DRM lists (Misirlisoy, 2004). It is also known that monitoring processes are more effective while using a foreign language compared to one's native language (Grant et al., 2023). These findings point to possible differences in the effectiveness of activation and monitoring processes for DRM and category lists

in a second language. This difference could be amplified by different levels of proficiency, as a lower level of proficiency is known to magnify the differences between a second and first language in many aspects (Suarez & Beato, 2021). Hence, an investigation of both paradigms while measuring proficiency in L2 to explore the possible outcomes would help clear the issue since would provide important implications and be a meaningful addition to the existing debate about the differences between the mechanisms DRM and Category lists adopt, as well as implications about the usage of second language and memory.

To our knowledge, the possible differential effects of proficiency in a second language on DRM and category lists has not yet been studied in the literature. The present study aims to investigate how proficiency in a second language will affect memory illusions that result from associative and semantic processes. The main research question is whether DRM and category lists will yield different results in terms of false and veridical memory levels for different levels of proficiency in a second language. As a result of the earlier mentioned difference in the levels of activation created by the 2 types of lists, and the proficiency levels in L2 amplifying this difference, an interaction effect of proficiency in L2 and list type is expected. As the association, and thus activation, increases with proficiency, false memory rate elicited by associative DRM lists is expected to be higher than those elicited by category lists. On the same note, for lower proficiency levels, since the association between words will not be as high as the high proficiency situation, they could be more susceptible to an illusion elicited by categorical relations rather than the more associative one created by DRM, leading to higher levels of false memory rates for category lists than DRM lists.

2.1 Method

This study was pre-registered in the Open Science Framework. All of the materials used in the study, as well as experiment scripts and data are available online at https://osf.io/qwgud/?view_only=9031cb8570c744d0a9d6338b9bba82e0

2.1.1 Participants

Since the effect size of interest was not previously reported by an earlier study in the literature, we designed our study to detect a moderate effect size of $f = .15$. An a-priori

power analysis using the G*Power-3 software (Faul et.al., 2007) indicated that with .80 power, to detect a moderate effect size f of .15 with alpha set at .05, a sample size of 76 would suffice. Because of possible loss of data due to reasons such as participants not completing the experiment or not having Turkish as their native language, we aimed to recruit a sample size of 10% more than the initial target number, which made 84 participants in total.

Participants were Middle East Technical University (METU) students who were 18–30-year-old native Turkish speakers. They had normal or corrected to normal visual acuity and did not report any problems with their sight and reading. Eighty-four participants participated in the experiment in exchange of course credit, a canteen coupon or voluntarily. Four participants were excluded from the analysis: 2 because of age restriction, 1 for failure to adhere with the experimental instructions, and 1 because of ambiguity in the dominant language. Analyses were carried out with the remaining 80 ($M = 22$, $SD=2.15$).

2.1.2 Apparatus and Stimuli

Stimulus presentation and data collection were carried out on the E-prime 2.0 software (Psychology Software Tools, Pittsburgh, PA) automatically.

2.1.2.1 Word Lists.

In both conditions, 15 lists were used, with 12 words on each list. Words on the lists were presented in descending order based on their level of associative strength relative to the critical word -for the DRM list- or their proportions of responses – for the category lists- and remained the same for each participant. However, the order of the lists was randomized for every participant.

Stimuli in the category list condition were selected from Van Overschelde et al.'s (2004) category norms (see Appendix A). For each category, typically, the first 15 words with the highest response proportions were used. However, the lists were monitored for words that could be problematic were eliminated or replaced with other words further down the category list. Specifically, words that might create confusion for non-native speakers, such as “cardinal” in the bird category that has more than one meaning; words that were also used in the DRM list condition, such as the word “leg”

which was in both the category list for body parts and the “table” list of DRM condition; words that consist of 2 words, such as cutting board in kitchen utensils category list; and words that would make the category more ambiguous, i.e. words “earthquake” and “monsoon” in the “weather phenomenon” category list were eliminated.

For the DRM condition, the lists were taken from the Stadler et al. (1999) study. Initially, lists with the highest proportion of false recognition for lure words were picked, while some lists with words that might not create the same levels of association for a non-native speaker, such as the list “black,” which encompassed words that have dual connotations, i.e. black and death. These lists were eliminated due to concerns about those figurative words not being seen as relevant to the rest of the list for a non-native speaker. After this selection, the lists went under the same monitoring process as the category lists, eliminating the words that were common with the lists in the category condition, that were too distinctive (e.g. longer than the rest of the words, less frequent), such as “stethoscope”; words that were not commonly known for non-native speakers, such as “swivel”; and proper nouns such as “Arctic” (see Appendix B).

For the recognition test of each condition, 60 words from the studied lists were selected. For each list, words in the 1st, 4th, 7th and 10th positions were selected to have equal difference in between the words, i.e. for the 12 words in the list, these positions allowed for equal spaces, with 3 words in between. For the non-studied items, there were 45 filler items (see Appendix C) and 15 critical lures, 1 critical lure for each list, 120 words total. The filler words were selected from MRC Psycholinguistics database that contains 150837 words (Coltheart, 1981). Forty-five words, each with 4 to 6 letters and no more than 3 syllables, were selected. The selected words had the following characteristics to standardize the filler words and make them less distinctive: concreteness scores between 400 and 600, imagery levels between 400 and 635, and frequency between 20 and 850.

2.1.2.2 English Proficiency Tests

At the end of the experiment, participants received placement tests to measure their proficiency levels, which consisted of a cloze test and a vocabulary test. Cloze tests are considered to be an efficient way to evaluate general language proficiency (Chae & Shin, 2015). In a cloze test, there is a reading passage with blanks, and the examinee

must realize relationships within or between the sentences in the passage and fill in the blanks accordingly (Chung & Ahn, 2019). The cloze test by Ionin and Montrul (2010), which has 40 multiple-choice questions, was used for our study.

For the vocabulary test, the updated version of the Vocabulary Levels Test (VLT) by Webb, Sasao, and Balance (2017) was used. This test aims to measure vocabulary knowledge at 5 different levels, with 30 words at each level.

Participants were also asked to report scores of EPE (English Proficiency Exam), which is the placement exam used by METU school of foreign languages to evaluate students' proficiency levels. EPE consists of 4 modules: Listening, reading, note-taking, and independent writing. It takes 165 minutes and is scored out of 100. This test aims to measure test-takers' ability to interpret and react to written and spoken academic materials, understanding interactions within classroom and campus environments, and generating written content (METU School of Foreign Languages, n.d.)

2.1.3 Design and Procedure

The study was approved by the METU Human Research Ethics Committee (see Appendix E). A 2 (list type: DRM, Category) x 3 (item type: target, CL, filler) within-subjects design was used to assess veridical and false memory rates for DRM and category lists. Language proficiency level was assessed as a continuous measure. The DRM and Category conditions were administered in different sessions. The order of the sessions was counterbalanced across participants.

Participants read and signed an informed consent form (see Appendix D) and answered demographic questions before starting the experiment. Their answers were collected automatically by the experiment program. The experiment consisted of 2 sessions- DRM and category- and each session had the same 3 phases: study phase, filler task, and the recognition test phase. Each session had 15 study lists with 12 words each, making 180 words total.

Before each phase of the experiment, participants were explained the instructions by the experimenter for the relevant phase, and they read the instructions on the screen as well. Then, they were asked to repeat the instructions to ensure there was no confusion about the task. After they summarized the instructions for the study phase, they started the experiment by pressing the spacebar key. The study phase started with a fixation

presented at the center of the screen for 2000 ms, followed by the stimuli, i.e. words. The words were presented in black at the center of the screen, on a light gray background for 1500 ms. with 1000 ms. inter stimulus intervals (see Figure 3). Due to the specific sequence of the words in the lists, i.e., descending in order of relevance to the CL or proportion of responses, the order of the words in each list remained the same and was not randomized so as not to undermine the memory illusion that it elicits. However, the order of lists was randomized for every participant. After every list, a “Next List” prompt was presented for 3000 ms. (see Figure 4). At the end of the study phase, participants were given a 4-minute filler task, requiring them to count down from 750 in increments of 7 out loud as the experimenter checked the answers and corrected when necessary. The participants completed the countdown task in Turkish. After the filler task was completed, the recognition test phase took place. In this phase, words were presented at the center of the screen in the same color, font, and background. Participants were asked to indicate whether they had seen the presented word in the study phase or not, by pressing the “yes” or “no” labeled keys, respectively. For these responses, the “d” and “k” keys on the keyboard were covered with a sticker indicating “yes” and “no” on them. The test phase was self-paced, with a maximum time of 5000 ms. If the participant did not respond within 5000 ms., the experiment program moved on to the next word. With the completion of the recognition test, the session ended.

There was a 5-minute break between the sessions, after which the experiment's second and last session was carried out. Both sessions consisted of the same 3 phases: the study phase, the filler task, and the recognition test phase. After the end of the second recognition test, participants received a brief survey that had questions about whether they realized any relationship between the words, whether they could keep their focus or got distracted, and finally, about their EPE scores.

Lastly, participants completed the proficiency tests. The total duration of the experiment was approximately 80 minutes, with approximately 25 minutes for each session, 5-minute break, and 25 minutes for the survey and the language proficiency tests, VLT, and the cloze test.

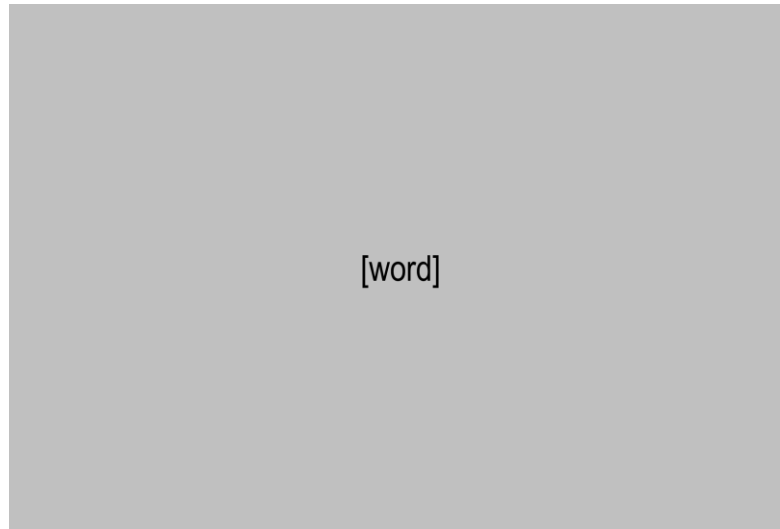
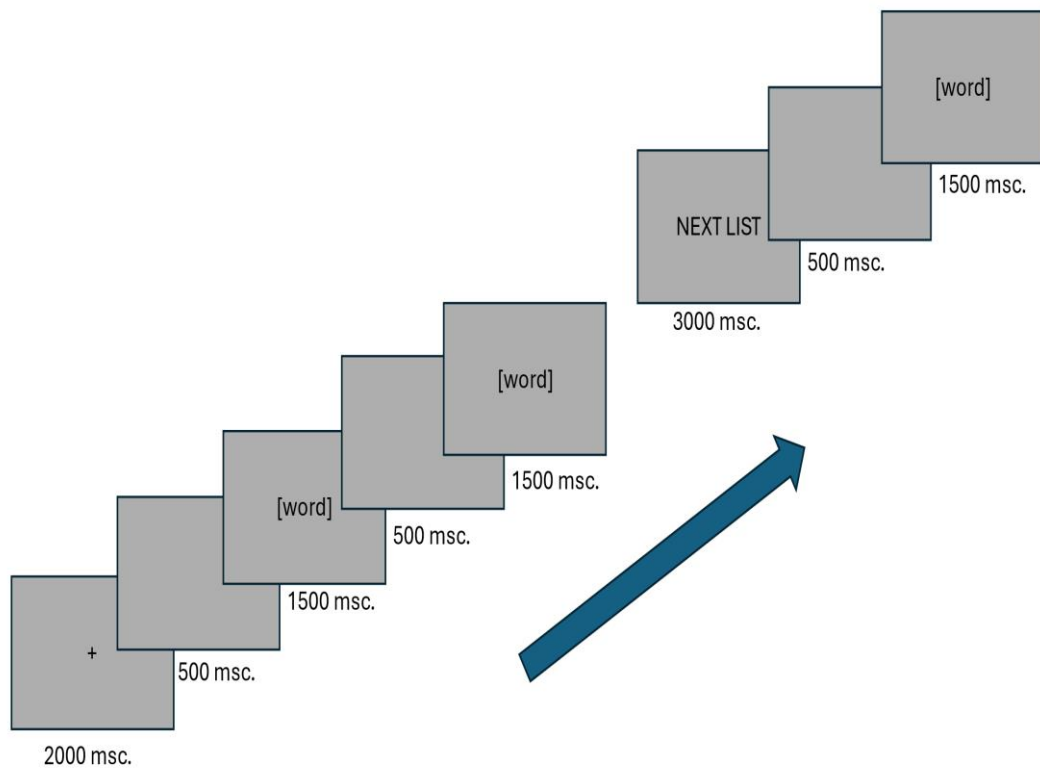


Figure 3. *Illustration of The Stimuli Presentation During The Study and The Test Phase*



Note. The phase starts with a fixation presented for 2000 ms, followed by a blank screen for 500 ms, after which the stimulus word is presented for 1500 ms. Each stimulus is followed by a 500 ms blank screen. Before the next stimulus is presented. After all 12 words in a list are presented, a “Next List” prompt is presented for 3000 ms, followed by a blank screen, and the trials go on until all 180 stimuli are presented.

Figure 4. *The Schematic Illustration of The Trial Sequence in The Study Phase*

CHAPTER 3

RESULTS

There were 84 participants in total, 62 of which were reported to be female, 21 were male, and 1 participant preferred not to say. Before the analysis, 4 participants' scores in the data were eliminated: 2 because of the age limit, 1 because of the dominant language not being clear, and 1 because of failure to adhere to the experiment procedure. The analyses were conducted with 80 participants. The mean age of the participants was 22 (21.97) and mean age of L2 acquisition was 10 (9.93).

3.1 List Type on Veridical and False Memory

Hit rates (HR), false alarm rates (FAR), and false memory rates (FMR) (false alarm for critical lures) were calculated by the proportion of "yes" responses to target (T), filler (F), and critical lure (CL) items. A 2 (list type: DRM, Category) x 3 (item type: target, CL, filler) within-subjects ANOVA on the proportion of "yes" responses was conducted to reveal the differential results of list type on hit and false memory rates.

There was a significant effect of List Type on the proportion of "yes" responses indicating that the frequency of participants' yes responses differed for DRM and category lists [$F(1, 79) = 32.416$, $MSE = .010$, $p < .001$, $\eta^2 = .291$]. The effect of Item Type on the proportion of "yes" responses was also significant [$F(2, 158) = 432.296$, $p < .001$, $MSE = .029$, $\eta^2 = .843$], indicating that participants' "yes" responses differed for target, filler and CL items.

The List Type and Item Type interaction was also significant [$F(2, 158) = 32.630$, $MSE = .010$, $p < .001$, $\eta^2 = .292$], suggesting that participants' "yes" responses for each type of item differed across list types.

Pairwise comparisons for the interaction between List Type and Item Type revealed that hit rates, indicated by the proportion of yes responses to target items did not

significantly differ for category lists (.729) and DRM lists (.693) (see Table 1), 95% CI [-.082, .011], $t(222) = -2.285$, Cohen's $d = -.237$, 95% CI [-.548, .075] $p_{\text{bonf}} = .348$ (Table 2). For the filler items, the proportion of "yes" responses was significantly higher for DRM lists (.183) compared to Category lists (.135), 95% CI [.002, .095], $t(222) = 3.108$, Cohen's $d = .322$, 95% CI [.006, .638], $p_{\text{bonf}} = .032$

Table 1 Means of Yes Response Rate Based on List and Item Type

Item Type	DRM	Category
Target (Hit Rate)	.693 (.015)	.729 (.015)
Critical Lure (FM Rate)	.575 (.02)	.432 (.021)
Filler (FA for filler items)	.183 (.014)	.135 (.014)

Note. Standard Errors are given in parentheses.

False memory rates, indicated by the proportion of yes responses to critical lures, also differed significantly between the list types. DRM lists elicited higher rates of false memory (.575) than category lists (.432) with a mean difference of .142. (95% CI [.096, .189], $t(222) = 9.099$, Cohen's $d = .944$, 95% CI [.565, 1.322], $p < .001$).

Table 2. Pairwise Comparisons of List and Item Type Levels

		MD	95% CI for MD			t	Cohen's d	95% CI for Cohen's d		p_{bonf}
			Lower	Upper	SE			Lower	Upper	
DRM, Target	Category, Target	-0.036	-0.082	0.011	0.016	-2.285	-0.237	-0.548	0.075	0.348
	DRM, Filler	0.51	0.444	0.575	0.022	23.051	3.377	2.472	4.281	< .001
	DRM, CL	0.118	0.053	0.184	0.022	5.35	0.784	0.313	1.254	< .001
Category-Target	Category, Filler	0.594	0.528	0.659	0.022	26.868	3.936	2.914	4.958	< .001
	Category, CL	0.296	0.231	0.362	0.022	13.408	1.964	1.331	2.597	< .001
DRM, Filler	Category, Filler	0.049	0.002	0.095	0.016	3.108	0.322	0.006	0.638	0.032
	DRM, CL	-0.391	-0.457	-0.326	0.022	-17.7	-2.593	-3.341	-1.845	< .001
Category Filler	Category, CL	-0.298	-0.363	-0.232	0.022	-13.46	-1.972	-2.606	-1.337	< .001

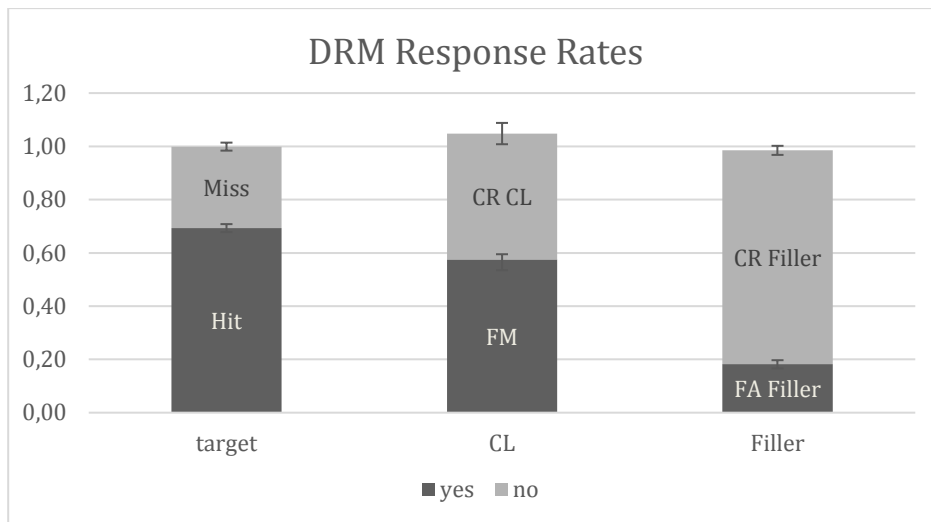


Figure 5. Mean Response Rates For The DRM Condition Across Item Types

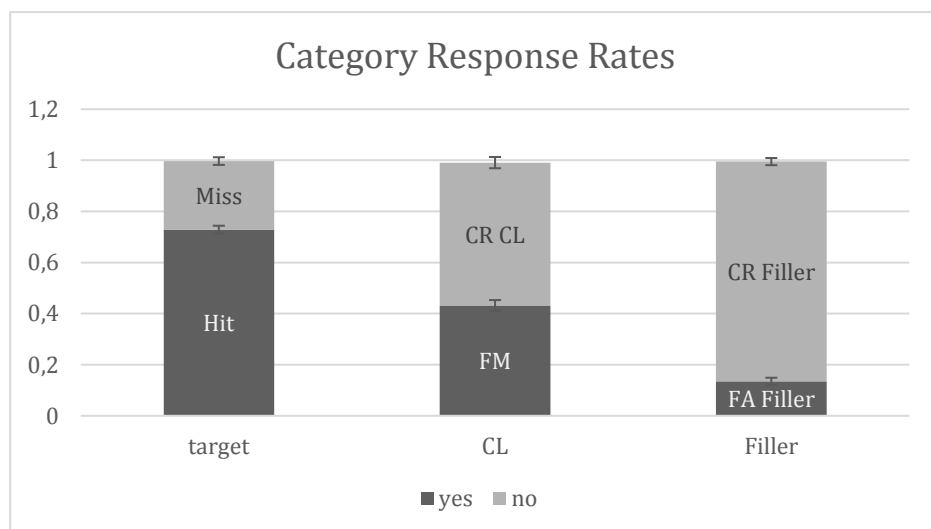


Figure 6. Mean Response Rates For The Category Condition Across Item Types

3.2 Proficiency, List Type, and False Memory

3.2.1 Proficiency Tests

To measure the L2 proficiency levels of participants, three different tests were used: EPE, VLT, and a Cloze test. EPE is the most comprehensive among the 3 tests, measuring proficiency, vocabulary, and grammar skills in many dimensions like

listening, reading, and writing. However, participants were administered EPE at different times, some of them years apart. Therefore, considering their current level of proficiency could be different than when they took the EPE, we provided 2 more tests in case EPE scores fail to capture participants' current level of proficiency, and a combined score was created by aggregating the scores in all 3 tests. The mean (M) and standard deviation (SD) values, range of participants' scores for each test and the possible range of scores can be seen in Table 2. We conducted separate analyses with both standardized EPE scores and the combined measure in case EPE scores alone fail to capture any possible effect of proficiency on false memory rates.

Table 3. *Descriptives For The 3 Proficiency Tests*

Descriptive Statistics			
	EPE	Cloze	VLT
Mean	74.946	32.333	131.048
Std. Deviation	12.585	3.399	16.891
Variance	158.374	11.55	285.299
Range	59.5	19	76
Minimum	35	23	74
Maximum	94.5	42	150

A correlation analysis was conducted to observe how participants' scores on these three different tests relate. Because the maximum possible score is different for each test (100 for EPE, 150 for VLT, and 40 for the Cloze test), z-scores for the values for each test were calculated.

The correlation analyses revealed a moderate correlation between EPE and Cloze test scores (Pearson correlation coefficient = 0.62), and strong positive correlations between EPE and VLT scores (0.75), EPE and combined scores (0.88), Cloze test and combined scores (0.86), VLT and combined scores (0.93) and Cloze test and LT.74).e

strong correlation between all the 3 tests and the combined measure indicates that the tests are consistent with each other (see Table 4).

Table 4. *Correlations Table For The 3 Proficiency Tests*

Pearson's Correlations					
Variable		EPE	Cloze	VLT	Combined
Cloze	Pearson's r	0.616	—		
	p-value	< .001	—		
VLT	Pearson's r	0.747	0.743	—	
	p-value	< .001	< .001	—	
Combined	Pearson's r	0.882	0.86	0.925	—
	p-value	< .001	< .001	< .001	—

Table 5. *Reliability Analysis For The 3 Proficiency Tests and Their Combined Scores*

Scale Reliability Statistics	
	Cronbach's α
Proficiency	0.937

Table 6. *Item Reliability Statistics For The 3 Proficiency Tests*

Item Reliability Statistics		
	Item-rest correlation	If item dropped Cronbach's α
EPE_zscore	0.786	0.939
ClozeTest_zscore	0.778	0.942
VLT_zscore	0.862	0.915

A reliability analysis was conducted to assess the internal consistency of the scales used to measure proficiency using Cronbach's Alpha. Because the maximum possible score is different for each test (100 for EPE, 150 for VLT, and 40 for the Cloze test), z-scores for the values for each test were calculated to standardize the scores and allow for comparison.

The reliability analysis revealed a high Cronbach's α (.937), indicating a high internal consistency among the 3 tests used (see Table 5). The impact of each item to the overall scale was revealed by item specific analysis, which showed that all items contributed positively to the internal consistency of the scale, with VLT (item-rest correlation=.862) and combined scores (item-rest correlation=1.00) being particularly important (see Table 6).

3.2.2 L2 Proficiency and List Type Interaction on False Memory

A multiple regression analysis which include proficiency scores, list type, and the interaction term on false memory rates were conducted to reveal the relationship between these variables. Because the focus of the study is on how proficiency interacts with list type to affect false memory rates, we conducted a moderated linear regression. In the first regression analysis, the EPE scores were used as the measure of proficiency, for that it is the most comprehensive of all 3 proficiency tests we utilized.

The scores for the EPE were centered to reduce multicollinearity by subtracting the mean score value from each data point. Because List type is a categorical variable, a dichotomous dummy variable for list type was used in the analyses. The regression equation with the mentioned variables is as follows:

$$\text{False Memory Rates} = \beta_0 + \beta_1(\text{Proficiency Level}) + \beta_2(\text{List Type}) + \beta_3(\text{Proficiency Level} \times \text{List Type}) + \epsilon$$

A linear regression analysis was conducted to examine the impact of proficiency (measured by centered EPE scores), list type (dummy coded as 1 for DRM vs. 0 for Category), and their interaction on false memory rates (FM Rate). The analysis revealed that the overall model was statistically significant, $F(3,156)=7.961$, $p<.001$, indicating that the predictors explained a significant portion of the variance in FM Rate. The full regression model explained 13.3% of the variance in FM Rate ($R^2=.133$, Adjusted $R^2=.116$ $RMSE=.186$).

The main effect of proficiency on FM rates was non-significant, indicating that proficiency levels do not have a statistically significant direct effect on false memory rates ($\beta=.0004$, $SE=.002$, $t(156)=.269$, $p=.788$), indicating that changes in proficiency levels alone do not explain variations in false memory rates.

The coefficient for list type was significant, suggesting that the type of list significantly affects false memory rates ($\beta = 0.142$, $SE = .029$, $t(156) = 4828$, $p < .001$).

The observed positive relationship with FM rates indicates that DRM lists are associated with higher false memory rates than category lists, as the ANOVA results also suggested (Table 3). Importantly, the interaction between proficiency levels and list type was not significant ($\beta = .0007$, $SE = .002$, $t(156) = .308$, $p = .758$), suggesting that the effect of list type on FM Rate does not significantly vary with different levels of proficiency.

Table 7. Mean and SE Values For Coefficients in The Model

	Descriptives			
	N	Mean	SD	SE
FM Rate	160	0.504	0.198	0.016
EPE Centered	160	-0.004	12.721	1.006
dummylisttype	160	0.500	0.502	0.040

Another regression analysis with the same variables was also conducted with the combined scores of all 3 proficiency tests instead of centered EPE scores. Despite being the most comprehensive among all the 3 proficiency tests we collected scores for, EPE was administered at different times for each participant, which may render its results less reflective of participants' current proficiency levels. To overcome this, the scores of all 3 tests were combined to have a more current and holistic measure of proficiency and used in the regression as the proficiency measure. The full regression model explained 13.4% of the variance in FM Rate ($R^2 = .134$ Adjusted $R^2 = .117$) The Root Mean Square Error (RMSE) of the model was .186. The main effect of L2 proficiency on false memory rates was not significant ($\beta = .003$, $SE = .023$, $t(156) = .126$, $p = .9$). The interaction was also not significant ($\beta = .017$, $SE = .033$, $t(156) = .515$, $p = .608$).

Overall, the results suggest that while list type has a notable effect on FM Rate, proficiency alone and its interaction with list type do not significantly impact false memory rates. The model's explanatory power, though statistically significant, accounted for a modest proportion of the variance in FM Rate.

3.3 Proficiency and List Type on Sensivity Index (d')

A signal detection analysis was conducted with hit, miss, false alarm and correct rejection rates. Signal detection theory is a framework used to understand how individuals make decisions to distinguish a “signal,” i.e., a studied item that should be recognized, from a “noise,” a non-target item, (d') and how their biases affect their decisions (c). It is widely used assess recognition performance (Banks, 1970).

Table 8. *Regression Coefficients For Predictors and Their Interaction on FM Rates*

Coefficients

Model		Unstd.	SE	Std.	t	p	95% CI	
							Lower	Upper
H ₀	(Intercept)	0.504	0.016		32.111	< .001	0.473	0.535
H ₁	(Intercept)	0.432	0.021		20.737	< .001	0.391	0.474
	EPE_Centered	4.420×10 ⁻⁴	0.002	0.028	0.269	0.788	-0.003	0.004
	List Type	0.142	0.029	0.360	4.828	< .001	0.084	0.201
	EPE_Centered * List Type	7.164×10 ⁻⁴	0.002	0.032	0.308	0.758	-0.004	0.005

Response bias refers to the tendency of participants to respond in a particular way, i.e. liberal or conservative. The threshold at which a participant decides whether an item is “old” (signal) or new (noise) is called the decision criterion (c). When participants adopt a liberal bias, or a liberal decision criterion (low c), they tend to give “yes” responses to the presented stimuli more often, which leads not only to higher hit rates (correctly identifying a signal), but also higher false alarm rates (incorrectly identifying a noise as a signal). On the other hand, a conservative bias (high c) lets participants to give “no” responses more often, which leads to more correct rejections (correctly rejecting a noise), as well as more misses (incorrectly rejecting a signal). Participants’ ability to distinguish between signal and noise is measured with d' (d-prime) sensitivity index, calculated by subtracting the z scores of false alarm rates from z score of hits rates. A higher d' indicates better discrimination, meaning the

participant can more effectively distinguish between items that were actually presented and those that were not. As mentioned earlier, the frequency of participants' yes responses differed for DRM and category lists. DRM lists elicited more "yes" responses for both target and critical lure items. This could point to a possible difference in participants' response strategies.

Table 9. *Model Summary and ANOVA Tables For d' analysis*

Model Summary - d'				
Model	R	R ²	Adjusted R ²	RMSE
H ₀	0.000	0.000	0.000	0.637
H ₁	0.389	0.152	0.135	0.592

ANOVA						
Model		Sum of Squares	df	Mean Square	F	p
H ₁	Regression	9.774	3	3.258	9.288	< .001
	Residual	54.725	156	0.351		
	Total	64.499	159			

Note. The intercept model is omitted, as no meaningful information can be shown.

To investigate this, regression analyses with proficiency (measured with combined scores) and list type on d' and c levels were conducted. The results of the regression analysis on d' levels revealed that the model was significant [$F(3,156)= 9.288$, $MSE= 3.258$, $p<.001$]. The full model explained 15.2% of the variance in d' levels ($R= .389$, $R^2= .152$, $RMSE= .592$).

Table 10. *Coefficients For The Covariates and The Model*

Coefficients								
Model	I	Unstandardize	Standar	Standardize	t	p	95% CI	
							d	d Error
H ₀	(Intercept)	1.347	0.05		26.76	< .001	1.248	1.447
H ₁	(Intercept)	1.539	0.066		23.24	< .001	1.409	1.67
	Proficiency	0.215	0.074	0.303	2.906	0.004	0.069	0.362
	dummylisttype	-0.384	0.094	-0.302	-	< .001	-	-
	Proficiency * dummylisttype	-0.095	0.105	-0.095	-0.91	0.364	0.302	0.112

Proficiency as measured by combined scores significantly predicted d' levels ($\beta=.215$, $SE=.074$, $t(156)=2.906$, 95% CI [.069, .362], $p=.004$). d' levels increased as proficiency scores increased. List type significantly predicted d' , ($\beta=-.384$, $SE=.094$, $t(156)=-4.098$, 95% CI [.569, .199], $p<.001$). More specifically, DRM condition was associated with lower levels of d' compared to category lists.

The interaction term between proficiency and list type on d' was not significant (see Table 10). These findings suggest that as the proficiency in L2 increases, memory sensitivity also increases. The main effect of list type suggested that participants had lower sensitivity in DRM list type condition than category list condition, yet proficiency did not affect this relationship significantly (see Figure 7).

The overall model did not significantly predict c levels, $F(3,156)= 1.673$, $MSE= .128$, $p=.175$, with an R^2 of .031. This suggests that the model explained 3.1% of the variance in c levels, but this effect was not statistically significant.

Table 11. *Model Summary and ANOVA Table For c Analysis*

Model Summary				
Model	R	R^2	Adjusted R^2	RMSE
H ₀	0.000	0.000	0.000	0.278
H ₁	0.177	0.031	0.013	0.277

ANOVA						
Model		Sum of Squares	df	Mean Square	F	p
H ₁	Regression	0.384	3	0.128	1.673	0.175
	Residual	11.937	156	0.077		
	Total	12.321	159			

Note. The intercept model is omitted, as no meaningful information can be shown.

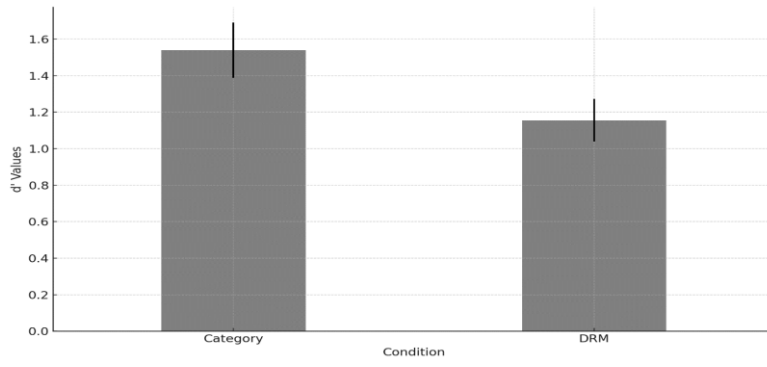


Figure 7. *d'* levels as a Recognition Sensitivity Measure Based on List Type.

CHAPTER 4

DISCUSSION

The current study aimed to investigate the effect of proficiency in a second language on false memory rates elicited by DRM and category lists. We expected an interaction between list types and proficiency on false memory rates. More specifically, we expected to find higher rates of DRM-elicited false memories compared to category-elicited false memory rates as proficiency scores increase, whereas, as proficiency scores decrease, we expected higher false memory rates in the category condition compared to DRM.

The results revealed a main effect of list type on the “yes” response rates for target, filler, and CL items, or, to put differently, rates of hit, FA for fillers, and false memory, respectively. DRM lists elicited higher “yes” response rates to CL and filler items. Hit rates did not differ for the list types. In other words, DRM lists elicited higher levels of false alarms than category lists. This could suggest a potential difference in the decision criteria of the participants for the two list types.

An analysis of d' levels as a measure of recognition sensitivity also revealed a significant difference between the two list types. DRM lists elicited lower levels of d' than category lists; indicating that participants could discriminate between targets and non-target items better in category condition. In addition to that, an analysis of c levels revealed no significant difference between the list types in terms of c levels, which indicates that DRM and category lists did not differ in terms of decision criterion. In other words, participants were not more liberal or conservative in either list type conditions. Thus, despite DRM lists have higher false alarm to filler item rates than category lists, the difference observed between the lists cannot be explained by a more liberal response bias. These result supports the finding that it is harder for participants to discriminate between the critical lure and studied items, leading to higher levels of false memory compared to category lists.

Many studies in the literature that investigate the difference between DRM and category list comparison reveal that DRM lists elicit higher levels of false recognition rates (Coane et al., 2021). Our results showed that this typical finding for DRM and category lists persists in participants' second language, too. To our knowledge, no study in the literature compares DRM and category lists in a second language, making this finding a valuable contribution. In addition to that, the false memory rate for DRM was lower than what is usually reported in the literature for DRM lists in the participants' native language. This result is parallel with the literature on that false memory levels elicited by DRM lists in L2 are lower than in native language. Moreover, the category list false memory rate (.43) was not much lower than the typical range reported in the literature for category lists (Dewhurst et al., 2009; Coane et al., 2016). This implies that L2 can, indeed, have differential effects on DRM and category-elicited false memories.

To investigate the effect of proficiency on false memory rates, we conducted 2 separate analyses with different measures for proficiency. In the first analysis, centered EPE results, which is the most comprehensive test among the 3 proficiency tests from which we gathered scores, were utilized as a measure of proficiency. The results revealed no significant effect of proficiency on false memory rates, suggesting that proficiency levels in L2 does not have a direct effect on false memory rates. Moreover, d' levels increased with proficiency scores, indicating that participants with higher proficiency levels could discriminate between target and non-target words better. This result is not in line with the findings in the literature. Studies investigating how proficiency levels affect false memories have many differences such as sample size, the native and dominant language of the participants, or how proficiency is measured (Suarez & Beato, 2021). These differences make comparison of the results harder. However, despite such differences, many of the studies on the topic agree that the DRM paradigm conducted in the first or dominant language of the participants elicits higher false memory rates compared to the second language (Kawasaki-Miyaji, 2003; Anastasi et al., 2005; Howe et al., 2008). Similarly, many studies showed that as well as native language, a higher proficiency in the second language also produces higher levels of false memory compared to lower proficiency (Beato & Arndt, 2021). Both of these observations can be explained by higher activation levels due to stronger connections between concepts due to more experience in a language (Howe et al., 2008). For

individuals with lower levels of proficiency in a second language, their network of concepts would be less interconnected, leading to the critical lure not being activated as strongly, making them less susceptible to false memory.

On the other hand, in studies where participants used both L1 and L2 frequently in their daily lives or where participants reported having academic education in the L2, the false memory levels between the two languages were similar. For example, a study done with Canadian participants who used both English and French daily found no difference in the false memory rates between the languages (Cabeza & Lennartson, 2005). This result could imply that for similar proficiency levels, frequent use of and experience in L2 could lessen the expected effect of proficiency on false memories. Considering that the participants in our study were METU students pursuing higher education in English and using their L2 daily in an academic context, their active use of the language might have affected the potential difference proficiency levels could create on false memory levels due to increased activity, and this could account for why our results differed from the previous findings.

There was also no significant effect of the list type and language proficiency interaction, which was the main focus of our study. The results imply that proficiency did not affect the differential false memory rates elicited by DRM and category lists. This result could be explained by treating proficiency as a continuous score instead of grouping participants as low and high proficiency levels. Despite using 3 different proficiency measures, i.e. EPE, VLT, and Cloze test, to capture possible differences in participants' proficiency levels, the scores on any of the tests did not have sufficient variance to imply clear differences in proficiency levels (see Table 2 for the range, variance, mean and *SD* values for each test). The scores did not exhibit substantial proficiency differences; therefore, a potential effect of proficiency levels on false memory rates could be missed simply due to not having different levels of proficiency. Studies in literature that report an effect of proficiency on false memory rates have different measures of proficiency, most of which are self-report (Suarez & Beato, 2021).

4.1 Limitations and Future Directions

As mentioned above, despite implementing 3 different assessments to address any potential variability in participants' proficiency levels, there was insufficient variance

in participants' proficiency levels to impact the false memory rates directly or to create a difference in how list type affects false memory rates. In future studies, it could be beneficial to ensure that participants have more pronounced differences in proficiency levels. Increasing the sample size would be another enhancement to provide better statistical power to ensure the detection of any existing effects.

4.2 Conclusion

Overall, the findings indicate that the differential effect of DRM and category lists is present in L2 as well, and not confined to native language. To our knowledge, the current study is the first to investigate the differential effects of DRM and category lists in a second-language context. Furthermore, the absence of significant findings regarding the effect of proficiency on false memory rates and the lack of an interaction effect with list type contrasts with existing literature that suggests higher proficiency in a second language enhances false memory rates, as well as creating a discrepancy with theoretical expectations, which propose that increased proficiency leads to higher activation levels, suggests that there may be an undetected effect of proficiency that our current study did not capture. All in all, the implications of these results are important for understanding how false memory mechanisms operate across L1 and L2 mechanisms and underscore the need for further research to explore this phenomenon in greater depth.

REFERENCES

- Anastasi, J. S., Rhodes, M. G., Marquez, S., & Velino, V. (2005). The incidence of false memories in native and non-native speakers. *Memory (Hove, England)*, *13*(8), 815–828. <https://doi.org/10.1080/09658210444000421>
- Arndt, J., & Beato, M. S. (2017). The role of language proficiency in producing false memories. *Journal of Memory and Language*, *95*, 146–158. <https://doi.org/10.1016/j.jml.2017.03.004>
- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, *74*(2), 81–99. <https://doi.org/10.1037/h0029531>
- Beato, M. S., Suarez, M., Cadavid, S., & Albuquerque, P. B. (2023). False memory in between-language conditions: A brief review on the effect of encoding and retrieving in different languages. *Frontiers in Psychology*, *14*. <https://doi.org/10.3389/fpsyg.2023.1237471>
- Bialystok, E., Dey, A., Sullivan, M. D., & Sommers, M. S. (2020). Using the DRM paradigm to assess language processing in monolinguals and bilinguals. *Memory & Cognition*, *48*(5), 870–883. <https://doi.org/10.3758/s13421-020-01016-6>
- Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164–169. <https://doi.org/10.1111/1467-8721.00192>
- Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review*, *110*(4), 762–784. <https://doi.org/10.1037/0033-295X.110.4.762>
- Buchanan, L., Brown, N. R., Cabeza, R., & Maitson, C. (1999). False memories and semantic lexicon arrangement. *Brain and Language*, *68*(1-2), 172–177. <https://doi.org/10.1006/brln.1999.2072>
- Chae, Eun-Young & Shin, Jeong-Ah. (2015). A study of a timed cloze test for evaluating L2 proficiency. *English Teaching*, *70*(3), 117-135.

- Chung, E. S., & Ahn, S. (2019). Examining Cloze Tests as a Measure of Linguistic Complexity in L2 Writing. *Language Research*, 55(3), 627–649. <https://doi.org/10.30961/lr.2019.55.3.627>
- Coane, J. H., & McBride, D. M. (2006). The role of test structure in creating false memories. *Memory & Cognition*, 34(5), 1026–1036. <https://doi.org/10.3758/BF03193249>
- Coane, J. H., McBride, D. M., Termonen, M. L., & Cutting, J. C. (2016). Categorical and associative relations increase false memory relative to purely associative relations. *Memory & Cognition*, 44(1), 37–49. <https://doi.org/10.3758/s13421-015-0543-1>
- Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 33A(4), 497–505. <https://doi.org/10.1080/14640748108400805>
- Deese J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58(1), 17–22. <https://doi.org/10.1037/h0046671>
- DeSoto, K. A., & Roediger, H. L. (2014). Positive and Negative Correlations Between Confidence and Accuracy for the Same Events in Recognition of Categorized Lists. *Psychological Science*, 25(3), 781–788. <https://doi.org/10.1177/0956797613516149>
- Dewhurst, S. A. (2001). Category repetition and false recognition: Effects of instance frequency and category size. *Journal of Memory and Language*, 44(1), 153–167. <https://doi.org/10.1006/jmla.2000.2738>
- Dewhurst, S. A. (2007). The effects of divided attention at study and test on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition*, 35(8), 1954–1965. <https://doi.org/10.3758/BF03192928>
- Dewhurst, S. A., & Robinson, C. A. (2004). False memories in children. Evidence for a shift from phonological to semantic associations. *Psychological Science*, 15(11), 782–786. <https://doi.org/10.1111/j.0956-7976.2004.00756.x>
- Dewhurst, S. A., Barry, C., & Holmes, S. J. (2005). Exploring the false recognition of category exemplars: Effects of divided attention and explicit generation.

European Journal of Cognitive Psychology, 17(6), 803–819.
<https://doi.org/10.1080/09541440540000013>

- Dewhurst, S. A., Bould, E., Knott, L. M., & Thorley, C. (2009). The roles of encoding and retrieval processes in associative and categorical memory illusions. *Journal of Memory and Language*, 60(1), 154-164. <https://doi.org/10.1016/j.jml.2008.09.002>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.
- Gallo, D. A., & Roediger, III, H. L. (2002a). Variability among word lists in eliciting memory illusions: Evidence for associative activation and monitoring. *Journal of Memory and Language*, 47(3), 469–497. [https://doi.org/10.1016/s0749-596x\(02\)00013-x](https://doi.org/10.1016/s0749-596x(02)00013-x)
- Goff, L. M., & Roediger, H. L. (1998). Imagination inflation for action events: Repeated imaginings lead to illusory recollections. *Memory & Cognition*, 26(1), 20–33. <https://doi.org/10.3758/bf03211367>
- Graves, D. F., & Altarriba, J. (2014). False memories in bilingual speakers. *Foundations of Bilingual Memory*, 205–221. https://doi.org/10.1007/978-1-4614-9218-4_10
- Howe, M. L., Wimmer, M. C., Gagnon, N., & Plumpton, S. (2009). An associative activation theory of children’s and adults’ memory illusions. *Journal of Memory and Language*, 60(2), 229–251. <http://dx.doi.org/10.1016/j.jml.2008.10.002>.
- Kawasaki-Miyaji, Y., T. Inoue and H. Yama. (2003). Cross-linguistic false recognition: How do Japanese-dominant bilinguals process two languages: Japanese and English? *Psychologia*, 46(4), 255–267. <https://doi.org/10.2117/psysoc.2003.255>
- Knott, L. M., & Dewhurst, S. A. (2007). The effects of divided attention at study and test on false recognition: A comparison of DRM and categorized lists. *Memory & Cognition*, 35(8), 1954–1965. <https://doi.org/10.3758/BF03192928>

- Knott, L. M., Howe, M. L., Toffalini, E., Shah, D., & Humphreys, L. (2018). The role of attention in immediate emotional false memory enhancement. *Emotion, 18*(8), 1063–1077. <https://doi.org/10.1037/emo0000407>
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language, 33*(2), 149–174. <https://doi.org/10.1006/jmla.1994.1008>
- Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research, 18*(2), 137–171.
- Marmolejo, G., Diliberto-Macaluso, K. A., & Altarriba, J. (2009). False memory in bilinguals: Does switching languages increase false memories? *The American Journal of Psychology, 122*(1), 1–16.
- Middle East Technical University School of Foreign Languages. *METU English Proficiency Exam (EPE)*. METU. (n.d.-a). Retrieved August 12, 2024, from https://epe.metu.edu.tr/index_en.htm
- Mısırlısoy, M. (2004). *Effects of associative processes on false memory: Evidence from converging associates and category associates procedures*. [Unpublished master's thesis]. Middle East Technical University.
- Oliveira, H. M., Albuquerque, P. B., & Saraiva, M. (2018). O Estudo das Falsas Memórias: Reflexão Histórica. *Temas Em Psicologia, 26*(4), 1763–1773. <https://doi.org/10.9788/tp2018.4-03pt>
- Pierce, B.H., Gallo, D.A., Weiss, J.A. et al. The modality effect in false recognition: Evidence for test-based monitoring. *Memory & Cognition 33*, 1407–1413 (2005). <https://doi.org/10.3758/BF03193373>
- Robinson, K. J., & Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science, 8*(3), 231–237. <https://doi.org/10.1111/j.1467-9280.1997.tb00417.x>
- Roediger III, H. L. (1996). Memory illusions. *Journal of memory and Language, 35*(2), 76-100.

- Roediger III, H. L., Jacoby, J. D., & McDermott, K. B. (1996). Misinformation effects in recall: Creating false memories through repeated retrieval. *Journal of Memory and Language*, 35(2), 300–318. <https://doi.org/10.1006/jmla.1996.0017>
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 803–814. <https://doi.org/10.1037/0278-7393.21.4.803>
- Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, 8(3), 385–407. <https://doi.org/10.3758/bf03196177>
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 2.0). [Computer software and manual]. Pittsburgh, PA: Psychology Software Tools Inc.
- Seamon, J. G., Luo, C. R., Schlegel, S. E., Greene, S. E., & Goldenberg, A. B. (2000). False memory for categorized pictures and words: The category associates procedure for studying memory errors in children and adults. *Journal of Memory and Language*, 42(1), 120–146. <https://doi.org/10.1006/jmla.1999.2676>
- Smith, S. M., Gerkens, D. R., Pierce, B. H., & Choi, H. (2002). The roles of associative responses at study and semantically guided recollection at test in false memory: The Kirkpatrick and Deese Hypotheses. *Journal of Memory and Language*, 47(3), 436–447. [https://doi.org/10.1016/s0749-596x\(02\)00012-8](https://doi.org/10.1016/s0749-596x(02)00012-8)
- Smith, S. M., Ward, T. B., Tindell, D. R., Sifonis, C. M., & Wilkenfeld, M. J. (2000). Category structure and created memories. *Memory & Cognition*, 28(3), 386–395. <https://doi.org/10.3758/bf03198554>
- Stadler, M. A., Roediger, H. L., 3rd, & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, 27(3), 494–500. <https://doi.org/10.3758/bf03211543>
- Suarez, M., & Beato, M. S. (2021). The role of language proficiency in False memory: A mini review. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.659434>

- Suarez, M., & Beato, M. S. (2023). False memory in a second language: The importance of controlling the knowledge of word meaning. *PloS one*, *18*(5), e0285747. <https://doi.org/10.1371/journal.pone.0285747>
- Sunderman, G., & Kroll, J. F. (2006). First language activation during Second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, *28*(03). <https://doi.org/10.1017/s0272263106060177>
- Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: effects of retention interval and level of processing. *Memory & cognition*, *29*(3), 424–432. <https://doi.org/10.3758/bf03196393>
- Tussing, A. A., & Greene, R. L. (1997). False recognition of associates: How robust is the effect? *Psychonomic Bulletin & Review*, *4*(4), 572–576. <https://doi.org/10.3758/bf03214351>
- Underwood, B. J. (1965a). False recognition produced by implicit verbal responses. *Journal of Experimental Psychology*, *70*(1), 122–129. <https://doi.org/10.1037/h0022014>
- Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*(3), 289–335. <https://doi.org/10.1016/j.jml.2003.10.003>
- Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics*, *168*(1), 34-70.
- Yüvrük, E., Turan, H., & Kapucu, A. (2019). Duygu içerikli Sözcüklerden Oluşan Türkçe DRM Listelerinin Geliştirilmesi. *Psikoloji Çalışmaları / Studies in Psychology*, 245–266. <https://doi.org/10.26650/sp2019-0026>

APPENDICES

A. CATEGORY WORD LISTS

List Number	1		2		3	
Category Name	Weather Phenomenon	RP	Relatives	RP	Metals	RP
CL	tornado	0.85	uncle	0.95	steel	0.62
1	hurricane	0.83	aunt	0.93	iron	0.57
2	rain	0.42	cousin	0.85	silver	0.55
3	snow	0.37	mom	0.8	copper	0.53
4	hail	0.24	grandma	0.79	gold	0.53
5	flood	0.21	father	0.77	aluminum	0.43
6	lightning	0.21	grandpa	0.74	platinum	0.17
7	blizzard	0.18	sister	0.74	titanium	0.17
8	thunder	0.14	brother	0.72	tin	0.15
9	wind	0.14	niece	0.32	bronze	0.14
10	storm	0.12	nephew	0.3	nickel	0.13
11	typhoon	0.11	son	0.08	lead	0.1
12	cloud	0.08	daughter	0.07	brass	0.09
Mean RP		0.3		0.62		0.32

List Number	4		5		6	
Category Name	4-Footed Animals	RP	Fabrics	RP	Colors	RP
CL	dog	0.98	silk	0.7	blue	1
1	cat	0.97	polyester	0.58	red	0.96
2	horse	0.52	wool	0.5	green	0.93
3	lion	0.41	nylon	0.26	yellow	0.92
4	bear	0.37	rayon	0.19	purple	0.83
5	tiger	0.36	satin	0.17	black	0.77
6	cow	0.35	denim	0.14	white	0.6
7	elephant	0.28	leather	0.13	pink	0.54
8	deer	0.23	linen	0.12	brown	0.4
9	mouse	0.23	velvet	0.1	gray	0.26
10	pig	0.21	suede	0.08	indigo	0.11
11	rat	0.19	cashmere	0.07	magenta	0.11

12	giraffe	0.16	lace	0.05	maroon	0.08
Mean RP		0.4		0.24		0.58

List Number	7		8		9	
Category Name	Kitchen Utensils	RP	Flowers	RP	Parts of Human Body	RP
CL	knife	0.95	rose	0.96	arm	0.82
1	fork	0.93	daisy	0.59	foot	0.71
2	spoon	0.93	tulip	0.53	finger	0.67
3	spatula	0.55	lily	0.3	head	0.61
4	pan	0.22	carnation	0.29	toe	0.61
5	pot	0.2	daffodil	0.23	hand	0.54
6	whisk	0.15	sunflower	0.23	ear	0.49
7	blender	0.14	dandelion	0.2	mouth	0.38
8	bowl	0.14	orchid	0.09	stomach	0.35
9	plate	0.13	petunia	0.09	knees	0.26
10	mixer	0.07	violet	0.09	neck	0.22
11	strainer	0.07	iris	0.08	elbow	0.17
12	stove	0.05	lilac	0.08	shoulder	0.17
Mean RP		0.35		0.29		0.46
List Number	10		11		12	
Category Name	Fruits	RP	Insects	RP	Vehicles	RP
CL	apple	0.95	fly	0.61	car	0.89
1	orange	0.86	ant	0.57	bus	0.58
2	banana	0.71	bee	0.41	truck	0.56
3	grape	0.52	mosquito	0.33	airplane	0.54
4	pear	0.5	beetle	0.31	train	0.44
5	peach	0.4	butterfly	0.19	bike	0.42
6	strawberry	0.4	wasp	0.17	van	0.32
7	kiwi	0.3	roach	0.16	motorcycle	0.3
8	pineapple	0.26	moth	0.15	subway	0.1
9	watermelon	0.24	centipede	0.9	taxi	0.1
10	plum	0.21	cricket	0.9	scooter	0.09
11	grapefruit	0.18	flea	0.5	helicopter	0.08
12	cherry	0.15	hornet	0.5	ship	0.04
Mean RP.		0.44		0.44		0.34

List Number	13		14		15	
Category Name	Vegetables	RP	Birds	RP	Precious Stones	RP
CL	carrot	0.77	eagle	0.47	diamond	0.92
1	lettuce	0.49	hawk	0.3	ruby	0.78
2	broccoli	0.42	crow	0.24	emerald	0.51
3	tomato	0.36	parrot	0.22	sapphire	0.49
4	cucumber	0.31	sparrow	0.21	pearl	0.22
5	peas	0.31	pigeon	0.2	gem	0.19
6	corn	0.3	seagull	0.13	amethyst	0.18
7	potato	0.28	dove	0.12	opal	0.18
8	celery	0.27	falcon	0.11	topaz	0.12
9	onion	0.24	Canary	0.1	garnet	0.11
10	spinach	0.14	Owl	0.1	jade	0.06
11	cauliflower	0.12	Penguin	0.09	quartz	0.06
12	cabbage	0.1	Raven	0.09	turquoise	0.05
Mean RP		0.32		0.18		0.3

B. DRM LISTS

List No	1	2	4	5
Old Proportion	0.84	0.84	0.82	0.81
CL	WINDOW	SMELL	CUP	SOFT
1	door	nose	mug	hard
2	glass	breathe	saucer	light
3	pane	sniff	tea	pillow
4	shade	aroma	lid	plush
5	ledge	hear	handle	loud
6	sill	see	coffee	cotton
7	house	nostril	straw	fur
8	open	scent	goblet	touch
9	curtain	reek	soup	fluffy
10	frame	stench	drink	feather
11	view	fragrance	plastic	skin
12	breeze	perfume	sip	tender
List No	6	7	8	9
Old Proportion	0.8	0.79	0.78	0.78
CL	SLEEP	ANGER	SWEET	TRASH
1	bed	mad	sour	garbage
2	rest	fear	candy	waste
3	awake	hate	sugar	can
4	tired	rage	bitter	refuse
5	dream	temper	good	sewage
6	wake	wrath	taste	bag
7	snooze	happy	tooth	junk
8	blanket	fight	nice	rubbish
9	doze	hatred	honey	sweep
10	slumber	mean	soda	scraps
11	snore	calm	chocolate	pile
12	nap	emotion	heart	dump

List No	10	11	12	13
Old Proportion	0.74	0.72	0.71	0.7
CL	CHAIR	HIGH	DOCTOR	THIEF
1	table	low	nurse	steal
2	sit	clouds	sick	robber
3	legs	up	lawyer	crook
4	seat	tall	medicine	burglar
5	couch	tower	health	money
6	desk	jump	hospital	cop
7	recliner	above	dentist	bad
8	sofa	building	physician	jail
9	wood	noon	ill	gun
10	cushion	cliff	patient	villain
11	stool	sky	surgeon	crime
12	bench	over	cure	bank
List No	14	15		
Old Proportion	0.69	0.64		
CL	MUSIC	BREAD		
1	note	butter		
2	sound	food		
3	piano	eat		
4	sing	sandwich		
5	radio	rye		
6	band	jam		
7	melody	milk		
8	horn	flour		
9	concert	jelly		
10	instrument	dough		
11	symphony	crust		
12	jazz	slice		

C. FILLER WORDS

DRM		CATEGORY		DRM		CATEGORY	
Word No	Word	Word No	Word	Word No	Word	Word No	Word
1	smile	1	Baby	25	cross	25	dinner
2	crowd	2	Post	26	fence	26	tape
3	card	3	Mine	27	gift	27	grass
4	phone	4	File	28	split	28	jury
5	pencil	5	Grave	29	beam	29	site
6	judge	6	Dust	30	vein	30	tail
7	trip	7	Object	31	savage	31	porch
8	track	8	blind	32	lock	32	atom
9	foam	9	symbol	33	craft	33	slave
10	mirror	10	speech	34	mark	34	novel
11	round	11	wheel	35	barrel	35	route
12	lunch	12	prison	36	team	36	sink
13	colony	13	task	37	base	37	king
14	widow	14	birth	38	circle	38	guest
15	mate	15	film	39	watch	39	text
16	meal	16	lane	40	signal	40	shop
17	mail	17	tree	41	park	41	youth
18	stone	18	bottle	42	match	42	autumn
19	palace	19	edge	43	dawn	43	sheet
20	weight	20	chart	44	shado w	44	penny
21	suit	21	lesson	45	poet	45	prize
22	goal	22	guide				
23	guard	23	victim				
24	faint	24	lord				

**D. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS
COMMITTEE**

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER



DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T. +90 312 210 22 91
F. +90 312 210 79 59
ueam@metu.edu.tr
www.ueam.metu.edu.tr

13 MART 2024

Konu: Değerlendirme Sonucu

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi: İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Mine Mısırlısoy Bıykoğlu

Danışmanlığınızı yürüttüğümüz Bahar Aykut'un "*İkinci Dilde Dil Yeterliliği Seviyelerinin Deese-Roediger-McDermott (DRM) Listeleri ve Kategorik Çağrışımlı Listeler ile Oluşturulan Bellek Yanılmalarına Etkisi*" başlıklı araştırmanız İnsan Araştırmaları Etik Kurulu tarafından uygun görülerek 0126-ODTÜİAEK-2024 protokol numarası ile onaylanmıştır

Bilgilerinize saygılarımla sunarım.

E. INFORMED CONSENT FORM

ARAŞTIRMAYA GÖNÜLLÜ KATILIM FORMU

Bu çalışma ODTÜ Psikoloji Bölümü yüksek lisans öğrencisi Bahar Aykut tarafından Prof. Dr. Mine Mısırlısoy danışmanlığında yürütülmektedir. Bu form sizi araştırma koşulları hakkında bilgilendirmek için hazırlanmıştır.

Çalışmanın Amacı Nedir?

Bu araştırma İngilizce kelime dağarcığı ve bellek ilişkisini araştırmaktadır.

Bize Nasıl Yardımcı Olmanızı İsteyeceğiz?

Çalışmada size bazı kelimeleri çalışmanız ve ardından kelimeleri tanımanız istenecektir. Araştırma yaklaşık 50 dk sürmektedir.

Sizden Topladığımız Bilgileri Nasıl Kullanacağız?

Araştırmaya katılımınız tamamen gönüllülük temelinde olmalıdır. Araştırmada sizden kimlik veya çalıştığınız kurum/bölüm/birim gibi belirleyici hiçbir bilgi istenmemektedir. Cevaplarımız tamamen gizli tutulacak, sadece araştırmacılar tarafından değerlendirilecektir. Katılımcılardan elde edilecek bilgiler toplu halde değerlendirilecek ve bilimsel yayınlarda kullanılacaktır. Sağladığımız veriler gönüllü katılım formlarında toplanan kimlik bilgileri ile eşleştirilmeyecektir.

Katılımla ilgili bilmeniz gerekenler:

Çalışma, genel olarak kişisel rahatsızlık verecek sorular içermemektedir. Ancak, katılım sırasında deneydeki sorulardan ya da herhangi başka bir nedenden ötürü kendinizi rahatsız hissederseniz cevaplama işini yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda çalışmayı uygulayan kişiye, çalışmadan çıkmak istediğinizi söylemek yeterli olacaktır.

Araştırmayla ilgili daha fazla bilgi almak isterseniz:

Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için Bahar Aykut (E-posta: bahar.aykut@metu.edu.tr) ya da Prof. Dr. Mine Mısırlısoy (mmine@metu.edu.tr) ile iletişim kurabilirsiniz.

*Yukarıdaki bilgileri okudum ve bu çalışmaya tamamen gönüllü olarak katılıyorum.
(Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).*

İsim Soyad

Tarih

İmza

---/---/---

F. TURKISH SUMMARY / TÜRKÇE ÖZET

BÖLÜM 1

GİRİŞ

Günlük yaşamımızda, anılarımız genellikle yaşadığımız olayların bir kaydı olarak görülür. Ancak, bellek sistemlerinin karmaşık doğası bazen hiç gerçekleşmemiş olayları hatırlamamıza veya gerçekleşmiş olayların detaylarını farklı anımsamamıza neden olabilir; bu fenomene yanlış bellek (Roediger & McDermott, 1995) denmektedir. Yanlış bellek üzerine yapılan araştırmalar, bu olguyu hem adli ve klinik açıdan, hem de bellek süreçlerini anlamak için bir araç olarak kullanılabilir bir bellek yanılsaması olarak incelemektedir. Erken dönem bellek araştırmaları çoğunlukla yanıt doğruluğuna odaklanmış ve hataları sadece performans sorunlarını gösteren işaretler olarak değerlendirmiştir. Ancak bazı çalışmalar bu hataların bellek mekanizmaları hakkında değerli bilgiler sağlayabileceğini belirtmiştir. Örneğin, Kirkpatrick'in 1894'teki çalışması, insanların çalıştıkları kelimelerle anlamsal açıdan ilişkili olan ancak kendilerine sunulmamış kelimeleri de çalıştıkları kelimelerden biriymiş gibi hatırlayabileceğini göstermiş ve önerilerin anıların nasıl raporlandığı üzerindeki etkisini vurgulamıştır.

Laboratuvar ortamında bellek yanılsamalarını araştırmak, yanlış anıların oluşumunu etkileyen faktörleri kontrol etmeyi ve manipüle edebilmeyi sağlar. Farklı uyarıcı türleri, kodlama ve geri çağırma teknikleri, ve etkilerini inceleyerek, bellek süreçlerinin aşamalarını daha iyi anlamak mümkündür. Yanlış bellek, bellek yanılsamaları olarak, mevcut bilgilerimizin belleğin yeni edililen bir bilgiyi nasıl organize ettiği ve geri çağırdığına olan etkisine dair ipuçları verir. Bu da, yukarıdan aşağıya (top-down) süreçlerin algımızı ve bellek işlemlerimizi nasıl şekillendirdiğine dair önemli bilgiler sağlar. Yanlış bellek deneysel ortamda pek çok yöntemle incelenebilmektedir. Bu yöntemlerden sıklıkla tercih edilen bir tanesi de, katılımcılara bir takım kelime listelerinin sunulduğu liste öğrenme paradigmasıdır. Liste öğrenme

paradigmasında, katılımcılar birbiriyle ilişkili kelimeler içeren kelime listelerini çalışmalarının ardından bir serbest geri çağırma veya tanıma testi gerçekleştirirler. Liste öğrenme paradigması yöntemlerinden en yaygın olarak kullanılanı Deese, Roediger ve McDermott (DRM) paradigmasıdır (Deese, 1959; Roediger & McDermott, 1995). Bu yöntem, Roediger ve McDermott tarafından (1995) Deese'in (1959) çalışmasının geliştirilmesi ile ortaya çıkmıştır. DRM paradigmasında katılımcılar, kelime listelerini çalışıp ardından bellek testine maruz kalırlar. Bu listelerdeki kelimeler, birbirleriyle ve listede yer almayan bir kritik kelime ile anlamsal ve çağrışımsal açıdan ilişkilidir (Gallo & Roediger, 2001). Daha sonra yapılan bellek testinde katılımcıların kendilerine çalışma aşamasında sunulmayan ancak çalıştıkları kelimelerle ilişkili olan bu kritik kelimeyi de sunulan kelimelerden biri olarak hatırlayıp hatırlamadığına bakılır. Örneğin, katılımcılara turta, kırmızı, armut gibi tamamı birbiriyle ve kritik kelime olan “elma” ile ilişkili kelimeler sunulduğunda, katılımcılar tipik olarak kritik kelime olan *elma* sözcüğünü de çalışmış gibi hatırlamaktadır. Araştırmalar, bu kritik kelimeler için olan yanlış bellek oranlarının genellikle gerçek çalışma öğelerinin serbest olarak hatırlanma veya tanınma oranlarına çok yakın olduğunu sonucuna varmıştır (Roediger & McDermott, 1995). Bu durum, DRM listelerinin sahte anılar oluşturma ve incelemede kullanılabilirlik önemli bir yöntem olduğunu göstermektedir.

DRM paradigması, bu yöntem kullanılarak oluşturulan yanlış belleğin farklı kodlama ve geri çağırma stratejileri (Dewhurst et al, 2017), yaş grupları (Carneiro et al., 2007), çalışma ve test aralıkları (Thapar & McDermott, 2001) kullanılan çalışmalarda da etkili ve güçlü olduğunun görülmesi sebebiyle güvenilir kabul edilmektedir. DRM listeleri kullanılarak oluşturulan yanlış belleğin genellikle gerçek anılardan ayırt etmesinin zor olduğu katılımcılar tarafından belirtilmiştir. Katılımcılar, kritik kelimeyi yüksek bir güvenle hatırlamaktadır (Dewhurst & Warrand, 2004). Benzer şekilde, katılımcılar sahte anılara sahte epizodik detaylar atfetmektedir (Payne et al., 1996).

Yanlış belleği incelemek için kullanılan bir diğer liste öğrenmesi yöntemi de Kategori İçi Çağrışımlar Yöntemidir (Buchanan et al., 1999), Bu yöntemde, kelime listeleri daha yüksek bir taksonomik kategorinin örneklerini içerir ve bu kategorinin en tipik örneği olan kelime kritik kelime olarak listeye dahil edilmez. Örneğin, meyve kategorisindeki bir liste armut, portakal ve muz gibi kelimeleri içerirken; meyve kategorisinin en tipik üyesi olan elma, kritik kelime olarak listeden çıkarılır. Ancak bu

kritik kelime, test aşamasında hala çalışılmış bir kelime olarak hatırlanabilir, bu da sahte bir anı yaratır. Kategori listeleri için sahte anı oranları genellikle DRM listelerinden daha düşüktür (Desoto & Roediger, 2014; Pierce et al., 2005).

DRM ve kategori listelerinin her ikisi de semantik ilişkileri kullanarak yanlış belleği tetiklese de, bu listeler içerdikleri kelimelerin birbirleri ve kritik kelimeyle ilişkileri açısından farklı doğalara sahiptir. Kategori listelerindeki kelimeler arasındaki anlam ilişkisi yalnızca taksonomik kategori üyeliği ile sınırlıyken, DRM listeleri birçok farklı çağrışım seviyesini içerir, eşanlamlılık, eşzamanlılık ve paylaşılan özellikler gibi (Smith et al., 2000; Coane et al., 2021). Örneğin, elma kritik kelimesine sahip bir kategori listesi armut, portakal ve muz gibi kelimeler içerebilirken, aynı elma kritik kelimesine sahip bir DRM listesi elma ile ilişkilendirilen turta (eşzamanlılık) ve kırmızı (sahip olunan özellik) gibi kelimeler içerebilir.

1.1.DRM ve Kategori Listelerinin Karşılaştırılması

DRM ve kategori listeleri arasındaki benzerlikleri ve farklılıkları incelemek, çeşitli teorik çıkarımlar geliştirerek yanlış belleğin nasıl oluştuğunu anlamamıza yardımcı olur. İki liste türünün benzerliklerini düşündüğümüzde, en önemli noktalardan birisi, hem DRM hem kategori listelerinin sunulmayan kelimelerin yanlış bellek oluşturmasını tetiklemek için semantik ilişkileri kullanmasıdır. Bu durum, iki yöntemin de bellek sistemlerinin ilişkili kavramları birbirine bağlama eğilimini kullandığını ve bu ilişkilendirme sonucunda kritik kelimeyi aktive ettiğine işaret etmektedir. Çalışmalar ayrıca her iki liste türünün de bir takım müdahalelerden benzer şekillerde etkilendiklerini göstermiştir. Örneğin, kodlama sırasındaki dikkat dağınıklığı veya uyarı gibi müdahaleler, her iki liste türü için yanlış bellek oranlarını azaltırken; listelerin bloklar halinde organize edilmesi ve listedeki öğe sayısının artırılması gibi faktörler her iki liste türü için de yanlış bellek oranlarını artırdığı bilinmektedir (Dewhurst & Anderson, 1999; Knott & Dewhurst, 2007). Bahsi geçen benzerliklerinin yanında, DRM ve kategori listelerinin önemli farkları bulunmaktadır. DRM ve kategori listelerinin yanlış bellek oluşumunda kullandıkları mekanizmaların birbirinden farklı olduğu düşünülmektedir (Coane et al., 2021). Kategori listeleri genellikle daha geniş bir kategorideki semantik ilişkileri kullanır; bu da listedeki öğelerin genel anlamlarına dayalı bir yanlış bellek oluşturulmasına yol açar. DRM

listeleri ise listedeki kelimeler arasındaki çağrışımsal bağlantıların gücünü kullanarak yanlış bellek oluşturur (Coane et al., 2021). DRM ve kategori listelerinin yanlış bellek oluşturmasındaki gözlemlenen bu farkların listelerin çağrışımsal doğasındaki farklardan ileri geldiği öne sürülmüştür (Knott et. al. 2012). DRM listelerinin genellikle daha yüksek oranda yanlış bellek üretmesinin sebebinin, DRM listelerinde daha yüksek oranda geriyedönük çağrışım gücüne (Backward Associative Strength - BAS) sahip olması olabileceği düşünülmüştür (Roediger et al., 2001). Geriyedönük çağrışım gücü, listede yer alan kelimelerden kritik kelimeye doğru olan çağrışımın gücünü ifade eder (Beato & Arndt, 2017). Yüksek BAS seviyelerine sahip olan DRM ve kategori listeleri arasında hem serbest geri getirme hem de tanıma görevlerinde oluşan sahte anı oranlarında önemli bir farkın olmadığı görülmüştür (Knott et. al., 2012). Ancak, Roediger ve arkadaşları (2001) ortalama BAS seviyesi daha düşük (.17) olan bir DRM listesinin, daha yüksek (.23) olan listeden daha yüksek oranda yanlış bellek üretebildiğini göstermiştir. Bu durum, listeler arasındaki sahte anı oranlarındaki farklılığın yalnızca BAS faktörü ile açıklanamayacağını ve farklı etkenlerin göz önünde bulundurulması gerektiğini ortaya koymuştur (Coane et.al., 2016). DRM ve kategori listeleri arasındaki bir başka farklılık da DRM listeleri tarafından oluşturulan yanlış bellek çalışma aşamasında oluşan çağrışımlardan etkilenmeye eğilimliken (Roediger et al., 2001) ve test aşamasındaki süreçlerden aynı oranda etkilenmezken (Coane & McBride, 2006), kategori listeleri tarafından oluşturulan yanlış belleğin test aşamasında meydana gelen semantik süreçlerden etkilenmesidir (Smith et al., 2002). DRM ve kategori liste türlerinin bir çok farklı olgu tarafından farklı şekillerde etkileniyor olmaları, kodlama ve geri çağırma süreçlerinde farklılıklar olabilmesi ve farklı seviyelerde yanlış bellek oranları üretiyor olmaları sebepleriyle, bu iki liste türünün farklı mekanizmalar kullandığı çıkarımı yapılmıştır (Smith et al., 2002). Dolayısıyla, bu iki liste türü arasında farklı sonuçlar oluşturabilecek faktörlerin araştırılması listelerin hangi etmenlerden ne şekilde farklı olarak etkilendiğini ve farklı mekanizmalar kullanıp kullanmadığının aydınlatması açısından faydalı olabilir.

1.2 Yanlış Belleği Açıklayan Yaklaşımlar

Yanlış bellek üzerine yapılan erken araştırmalar, zihnin eksik bilgileri nasıl telafi ettiğini incelemiş ve belleğin, olayların sadece hatırlanması değil, yeniden

yapılandırılmasında da rol oynadığını ortaya koydu. Bartlett'in (1932) çalışması, katılımcıların bir hikaye okuduktan sonra, orijinalinde olmayan detayları hatırladığını gösterdi. Bu durum, belleğin sadece olayların geri getirilmesi değil; aynı zamanda yeniden yapılandırıcı süreçleri içerdiğini önermektedir (Brainerd & Reyna, 2005). Yapılandırıcı teoriler, bireylerin bir olayın anlamını hatırlama eğiliminde olduklarını, ayrıntılardan ziyade genel anlamı hatırladıklarını savunur (Reyna, 1996). Bransford ve Franks (1971), insanların bir olaydan hemen sonra belirli detayları unuttuğunu ve kendi yorumlarına dayalı entegre bir temsil tutuklarını önerdiler.

Şema temelli yaklaşımlar, belleğin şemalar veya ilgili kavramların ağları şeklinde organize edildiğine odaklanır. Bu yaklaşım, bireyler bir dizi birbirine ilişkili kelimeyle karşılaştıklarında, bu kelimeler için ortak olan bir şemanın aktive olduğunu ve bunun, kritik kelimeyi de içeren ilişkili kavramlara yayıldığını önerir. Bu aktivasyon, listede olmayan kelimenin tanıdık görünmesini sağlar ve yanlış tanımaya katkıda bulunur. Kritik kelimenin ortak şemaya uygunluğu, yanlış belleğin oluşumunu destekler.

Yanlış belleğin nasıl oluştuğunu açıklayan en yaygın yaklaşımlardan biri Aktivasyon/İzleme teorisidir (Activation Monitoring Framework - AMF) (Roediger et al., 2001). Bu yaklaşıma göre listede bulunan semantik olarak ilişkili kelimeler çalışıldığında, bu kelimeler aktive olur. Bu aktivasyon, sunulmayan kritik kelimeye de yayılır. Bu aktivasyon miktarı arttıkça, çeşitli faktörler, örneğin Geriye Dönük İlişkilendirme Gücü (BAS), kelime sayısı ve kritik kelime sayısı gibi, kritik kelimenin çalışılmış listede yer almış gibi kabul edilmesi daha kolay hale gelir. Diğer taraftan, çağrışımın artması, bu aktivasyonun kaynağını izlemeyi zorlaştırır, bu nedenle öğeyi "yeni" olarak tanımlamak daha zor hale gelir. Maruz kalma ve yeterlilik arttıkça, kavramların ve bunların ilişkisel bağlantılarının otomatik olarak aktive edilmesi de artar; bu da yanlış bellek üretmenin bilinçli bir çaba veya farkındalık olmadan gerçekleşmesine neden olur (Howe, 2005).

Yanlış belleği açıklamak için kullanılan bir diğer teorik çerçeve Bulanık İzler Teorisidir (Fuzzy Trace Theory - FTT). Bu teori, bellekte detay özet izler olmak üzere iki tür iz oluştuğunu öne sürer. Detay izler, bir olayın fiziksel özellikleri gibi tam detaylarını ifade eder. Özet izleri ise olayın ana temasını, anlamını veya genel desenini içerir (Brainerd & Reyna, 2002). Örneğin, bir kelime listesi sunulduğunda, kelimenin

renge ve yazı tipi veya sunum zamanı gibi görsel özellikler detay izleri olurken, listenin genel teması ise özet izdir.

FTT, semantik olarak ilişkili kelimeler incelendiğinde, katılımcıların listenin genel temasının veya anlamının bir özetini, yani özet izini oluşturduğunu öne sürer. Katılımcılar, test kelimeleri hakkında eski veya yeni kararları verirken bu özet izine dayanırlar. Eğer kritik kelime özet izine yakınsa, bu kelimenin listenin bir parçası olarak yanlış hatırlanma olasılığı artar. Detay izlerin geri getirilmesi bu olasılığı sınırlayabilir.

1.4 İkinci Dilde Yanlış Bellek ve Yeterliliğin Rolü

Yanlış bellek üzerine yapılan araştırmalar, özellikle DRM listeleri kullanılarak yapılan çalışmalar, bu etkilerin çeşitli dillerde (Anastasi et al., 2005 İspanyolca; Nam, 2018, Korece; Misirlisoy, 2004, Türkçe) ortaya çıktığını göstermiştir. Ancak, dil becerileri, iki dilli olmanın tanımları ve liste öğelerindeki farklılıklar gibi faktörler nedeniyle farklı dillerde yapılan çalışmaların sonuçlarını karşılaştırmak zorlayıcı olabilmektedir (Beato ve arkadaşları, 2023). Buna rağmen, çalışmaların genelinde yanlış bellek kişinin baskın dilinde (L1) daha belirgin olduğu ve ikinci dilde (L2) yeterlilik arttıkça yanlış bellek oranlarının da yükseldiği genel bir bulgu olarak görülmektedir (Kawasaki-Miyaji ve arkadaşları., 2003; Sunderman & Kroll, 2006).

Ters Hiyerarşik Model (Reverse Hiearchical Model - RHM), düşük yeterlilikteki iki dilli bireylerin L2'yi anlamak için daha çok L1'e güvendiklerini, bu yüzden L2 kelimelerinin aktivasyonunun zayıf olduğunu ve dolayısıyla yanlış bellek oranlarının düşük olduğunu açıklar. Yeterlilik arttıkça, L2 bağlantıları güçlenir ve bu da daha yüksek aktivasyon ve daha fazla yanlış belleğe yol açar. Bu durum, L1 ve L2'nin farklı şekilde işlendiğini gösterir, ancak bu farkların nedenleri henüz tam olarak anlaşılmamıştır.

Yeterlilik üzerindeki etkileri farklı ilişkilendirme yöntemleri, örneğin DRM ve kategori listeleri, dil yeterliliğinin yanlış bellek oranlarını ve bu etkilerin arkasındaki mekanizmaları nasıl etkilediğini anlamak için önemli bilgiler sağlayabilir.

BÖLÜM 2

MEVCUT ÇALIŞMA

Önceki bölümde vurgulanan noktalar doğrultusunda, DRM ve kategori listeleri içerdikleri kelimelerin birbiriyle ve kritik kelimeyle olan çağrışımsal bağları açısından önemli farklılıklar gösterir ve bu farklılıklar iki liste türünün oluşturduğu yanlış belleğin farklı olmasında da etkilidir. Bu farklılıklar genellikle, her iki liste türündeki kelimeler arasındaki ilişkilerin doğası ve bu ilişkilerin ürettiği farklı aktivasyon seviyeleriyle ilişkilendirilir. DRM listeleri, tipik olarak, daha karmaşık çağrışımsal ilişkilere sahip olması nedeniyle kategori listelerinden daha yüksek oranda yanlış bellek üretir.

Mevcut araştırmalar, ikinci dil (L2) yeterliliği seviyelerinin artmasıyla yanlış bellek oranlarının arttığını göstermiştir (Arndt & Beato, 2017). Ayrıca, katılımcılar, genellikle L1 ile kıyaslandığında L2'deki fonolojik ve ortografik özelliklerin manipülasyonlarına daha duyarlıdır. Bu durum genellikle L1'deki kavramların zihinsel temsilleri ve aralarındaki çağrışım gücünün L2'ye kıyasla daha güçlü olması ve L2 yeterliliği arttıkça bu ilişkilerin güçlenmesi ile açıklanır (Dewhurst & Robinson, 2004). Bu durum, L2'deki zihinsel temsiller ve ilişkilerin farklılıklarının L1'e kıyasla farklı sahte anı sonuçlarına yol açabileceğini önermektedir (Arndt & Beato, 2017).

DRM listelerindeki daha yüksek yanlış bellek seviyeleri için bir diğer açıklama denetleme yaklaşımıyla ilgilidir. Bu teori, DRM listelerinin denetleme süreçlerinden daha az etkilendiğini, bunun aksine kategori listelerinin izleme süreçlerinden daha fazla etkilendiğini öne sürer (Misirlisoy, 2004). İzleme süreçlerinin L1'de L2'ye kıyasla daha etkili olduğu bilinmektedir (Grant et al., 2023). Bu, DRM ve kategori listeleri arasındaki aktivasyon ve denetleme süreçlerindeki farklılıkların ikinci dilde daha belirgin olabileceğini ve bu farklılıkların yeterlilik seviyeleri tarafından güçlendirilebileceğini önermektedir (Suarez & Beato, 2021).

Mevcut literatür ve DRM ile kategori listeleri arasındaki gözlemlenen farklılıklar dikkate alındığında, bu çalışma, L2 yeterliliğinin çağrışımsal ve anlamsal süreçlerden kaynaklanan bellek illüzyonları üzerindeki etkilerini incelemeyi amaçlamaktadır. Özellikle, DRM ve kategori listelerinin farklı L2 yeterlilik seviyelerinde yanlış bellek oranları açısından farklı sonuçlar üretip üretmediğini araştıracaktır.

Çalışmanın Önemi

Bu çalışma, L2 yeterliliğinin bellek süreçlerini, özellikle yanlış belleği nasıl etkilediğini anlamaya katkıda bulunacaktır. DRM ve kategori listelerinin farklı yeterlilik seviyelerinde karşılaştırılması, bu iki liste türü arasındaki tipik farklılıkların ikinci dilde devam edip etmediğini ve yeterlilik seviyelerinin bu etkileri nasıl değiştirdiğini ortaya koyacaktır. Ayrıca, bulgular ikinci dil edinimi ile bellek güvenilirliği üzerindeki pratik sonuçlar için önemli çıkarımlarda bulunacaktır. Bu araştırmanın DRM ve kategori listeleri için L2'deki yeterliliğin farklı etkilerini karşılaştıran ilk çalışma olması, hem bellek hem de ikinci dil işleme süreçleri açısından önemli sonuçlar doğuracaktır.

2.1 Yöntem

Bu çalışmanın ön kaydı Open Science Framework sitesine kaydedilmiştir. Çalışmada kullanılan tüm materyaller, deney senaryoları ve verilere https://osf.io/qwgud/?view_only=9031cb8570c744d0a9d6338b9bba82e0 adresinden ulaşılabilir.

2.2 Katılımcılar

G*Power-3 yazılımı (Faul et al., 2007) kullanılarak yapılan güç analizinde, 0.80 güç ve alfa seviyesinin 0.05 olarak belirlendiği durumda, orta düzeyde bir etki büyüklüğü $f = 0.15$ 'i tespit etmek için 76 kişilik bir örneklem büyüklüğünün yeterli olacağını belirlenmiştir. Deneyin tamamlanmaması veya Türkçe'nin ana dil olmaması gibi veri kaybı olasılıkları göz önünde bulundurularak, başlangıçta hedeflenen sayının %10 fazlası kadar katılımcı alınması düşünülerek toplam katılımcı sayısı 84 olarak belirlenmiştir.

18-30 yaş aralığında, anadili Türkçe olan Orta Doğu Teknik Üniversitesi (ODTÜ) öğrencileri deneye, ek puan veya kantin kuponu karşılığında ya da gönüllü olarak katılmıştır.

Toplamda 84 kişiden veri toplanmış, 2 kişi yaş sınırı, 1 kişi deney talimatlarına uymaması ve 1 kişi de ana dili konusundaki belirsizlik nedenlerinden olmak üzere toplamda 4 katılımcının verisi analize dahil edilmemiştir.. Analizler, geriye kalan 80 katılımcı (ortalama yaş: 22, SS= 2.15) ile yapılmıştır.

2.3 Veri Toplama Araçları ve Uyarılar

Uyarıcıların sunumu ve veri toplama, E-prime 2.0 yazılımı (Psychology Software Tools, Pittsburgh, PA) kullanılarak otomatik olarak gerçekleştirilmiştir.

2.3.1 Kelime Listeleri

Her iki koşulda da, her biri 12 kelimedenden oluşan toplamda 15 liste kullanılmıştır. Kategori listesi koşulundaki uyarıcılar, Van Overschelde ve arkadaşlarının (2004) kategori normlarından seçilmiştir (Bkz. Ek A).

Her kategori için genellikle en yüksek yanıt oranlarına sahip ilk 15 kelime kullanılmıştır. Ancak, sorun oluşturabileceği düşünülen kelimeler ya elenmiş ya da listeye daha gerideki kelimelerle değiştirilmiştir.

DRM koşulu için listeler, Stadler ve arkadaşlarının (1999) çalışmasından alınmıştır.

Her koşul için tanıma testinde, çalışılan listelerden 60 kelime seçilmiştir. Her liste için, 1., 4., 7. ve 10. pozisyondaki kelimeler seçilmiştir, böylece kelimeler arasında eşit mesafeler sağlanmıştır. Katılımcılar tarafından çalışılmamış olan kelimelerden 45'i çeldirici (Bkz. Ek C) 15'i ise kritik kelime olmak üzere 60 çalışılmayan kelime kullanılmıştır. Tanıma testi toplamda 120 kelime içermektedir.

2.3.2 İngilizce Yeterlilik Testleri

Deneyin sonunda, katılımcılara İngilizce yeterlilik seviyelerini ölçen testler verilmiştir. Bu testler, Ionin ve Montrul (2010) tarafından hazırlanmış 40 çoktan seçmeli sorudan oluşan boşluk doldurma testi ve Webb, Sasao ve Balance (2017) tarafından güncellenmiş Vocabulary Levels Test (VLT) kullanılmıştır. Katılımcılardan ayrıca METU Yabancı Diller Okulu tarafından öğrencilerin yeterlilik seviyelerini değerlendirmek için kullanılan EPE (İngilizce Yeterlilik Sınavı) puanlarını raporlamaları istenmiştir.

2.4 Araştırma Tasarımı ve İşlem

Araştırma, METU İnsan Araştırmaları Etik Komitesi tarafından onaylanmıştır (Bkz. Ek E). Doğru ve yanlış bellek oranlarını değerlendirmek için 2 (Liste türü: DRM,

Kategori) x 3 (Kelime türü: hedef, kritik kelime, çeldirici) denek içi tasarımı kullanılmıştır. Dil yeterlilik seviyesi sürekli bir ölçüm olarak değerlendirilmiştir. DRM ve Kategori koşulları farklı oturumlarda uygulanmıştır. Oturumların sırası katılımcılar arasında dengelenmiş olarak belirlenmiştir.

Katılımcılar, deney başlamadan önce bir bilgilendirilmiş onay formunu imzalamış (Bkz. Ek D) ve demografik soruları cevaplamıştır. Yanıtlar otomatik olarak deney programı tarafından toplanmıştır. Deney, her biri aynı 3 aşamadan oluşan 2 oturumdan oluşmaktadır: çalışma aşaması, çeldirici görevi ve tanıma testi aşaması. Her oturumda, her biri 12 kelime içeren 15 çalışma listesi bulunmakta, toplamda 180 kelime içermektedir. Kelimeler, ekranda ortada, açık gri bir arka planda 1500 ms süreyle ve 1000 ms aralıklı olarak siyah renkte gösterilmiştir. Her listenin ardından, 3000 ms süreyle "Sonraki Liste" istemi sunulmuştur. Çalışma aşamasının sonunda, katılımcılara 750'den 7'şer 7'şer saymaları gereken 4 dakikalık bir çeldirici görevi verilmiştir. Çeldirici görevin ardından tanıma testi aşaması gerçekleştirilmiştir. Bu aşamada, kelimeler ekranda aynı renk, yazı tipi ve arka planda gösterilmiştir. Katılımcılardan, sunulan kelimenin çalışma aşamasında görüp görmediklerini belirtmeleri istenmiştir, "evet" ya da "hayır" olarak etiketlenmiş tuşlara basarak yanıt vermişlerdir. Test aşaması katılımcının kendi hızında ilerlemiştir, maksimum süre 5000 ms'dir. Tanıma testi tamamlandığında, oturum sona ermiştir. Oturumlar arasında 5 dakikalık bir ara verilmiş ve ardından deneyin ikinci ve son oturumu gerçekleştirilmiştir. Her iki oturumda da çalışma aşaması, çeldirici görevi ve tanıma testi aşaması olmak üzere 3 aşama bulunmaktadır. İkinci tanıma testinin sonunda katılımcılar yeterlilik testlerini tamamlamıştır. Deneyin toplam süresi yaklaşık 80 dakika olup, her oturum için yaklaşık 25 dakika, 5 dakikalık bir ara ve anket ile dil yeterlilik testleri, VLT ve boşluk doldurma testi için 25 dakika sürmüştür.

BÖLÜM 3

BULGULAR

Her katılımcı için yanıtları ve yanıt tepki süreleri toplanmıştır. Kelime tanıma yanıtları kullanılarak, kelime tanıma isabet oranları, ve yanlış alarm oranları ve kritik kelimeye olan yanlış alarm oranı hesaplanmıştır. Tanıma performansını ölçmek

için verilen yanıtların duyarlılığını ölçümleyen d' parametresi isabet ve yanlış alarm oranları kullanılarak hesaplanmıştır. Liste türünün gerçek ve sahte bellek oranları üzerindeki farklı etkilerini ortaya koymak için, “evet” yanıtları üzerinde 2 (Liste türü: DRM, Kategori) x 3 (Kelime türü: hedef, KL, dolgu) tekrarlı ölçüm ANOVA analizi gerçekleştirilmiştir.

Liste Türünün "evet" yanıtlarının oranı üzerindeki etkisi anlamlı bulunmuştur, bu da DRM ve kategori listeleri için katılımcıların "evet" yanıtlarının sıklığının farklı olduğunu göstermektedir [$F(1, 79) = 32.416, MSE = .010, p < .001, \eta^2 = .291$]. Kelime türünün "evet" yanıtlarının oranı üzerindeki etkisi de anlamlı bulunmuştur ($F(2, 158) = 432.296, p < .001, MSE = .029, \eta^2 = .843$), bu da katılımcıların "evet" yanıtlarının hedef, çeldirici ve kritik kelimeler için farklı olduğunu göstermektedir.

Liste türü ve kelime türü etkileşimi de anlamlı bulunmuştur ($F(2, 158) = 32.630, MSE = .010, p < .001, \eta^2 = .292$), bu da katılımcıların her kelime türü için "evet" yanıtlarının liste türlerine göre farklılık gösterdiğini öne sürmektedir.

Liste Türü ve kelime Türü etkileşimi için yapılan çiftli karşılaştırmalar, hedef kelimelere verilen “evet” yanıtlarının oranı açısından kategori listeleri (.729) ile DRM listeleri (.693) arasında anlamlı bir fark bulunmadığını göstermiştir (Tablo 1), 95% CI [-.082, .011], $t(222) = -2.285$, Cohen's $d = -.237$, 95% CI [-.548, .075], $p_{\text{bonf}} = .348$. Çeldirici kelimeler için, “evet” yanıtlarının oranı DRM listelerinde (.183) Kategori listelerine (.135) göre anlamlı şekilde daha yüksek bulunmuştur, 95% CI [.002, .095], $t(222) = 3.108$, Cohen's $d = .322$, 95% CI [.006, .638], $p_{\text{bonf}} = .032$). Yanlış bellek oranları, yani kritik uyarıcılara verilen “evet” yanıtlarının oranı liste türleri arasında anlamlı bir şekilde farklılık göstermiştir. DRM listeleri, kategori listelerinden (.432) daha yüksek yanlış bellek oranlarına (.575) neden olmuştur ve ortalama fark .142'dir (95% CI [.096, .189], $t(222) = 9.099$, Cohen's $d = .944$, 95% CI [-.565, 1.322], $p < .001$).

Katılımcıların ikinci dil (L2) yeterlilik seviyelerini ölçmek için üç farklı test kullanılmıştır: EPE (English Proficiency Exam – İngilizce Yeterlilik Sınavı), VLT (Vocabulary Levels Test- Söz Dağarcığı Seviyeleri Testi) ve Cloze test. EPE, dinleme, okuma ve yazma gibi birçok boyutta yeterlilik, kelime bilgisi ve dil bilgisi becerilerini ölçen kapsamlı bir testtir. Ancak, bazı katılımcılara EPE testleri farklı zamanlarda uygulanmış, bu da mevcut yeterlilik seviyelerinin test zamanında olduğundan farklı olabileceğini göstermektedir. Bu nedenle, EPE puanlarının katılımcıların mevcut

yeterlilik seviyelerini yansıtmama ihtimaline karşı ek testler uygulanmıştır. Analizlerde hem standartlaştırılmış EPE puanları hem de tüm üç testin puanlarının birleştirilmesiyle oluşturulan bir birleşik skor kullanılmıştır.

Katılımcıların bu üç farklı testteki puanlarının ilişkisini gözlemlemek için bir korelasyon analizi yapılmıştır. Her testin maksimum puanı farklı olduğundan (EPE için 100, VLT için 150, ve Cloze testi için 40), her testin z-puanları hesaplanmıştır. Korelasyon analizleri, EPE ve Cloze testi puanları arasında orta düzeyde bir korelasyon (Pearson korelasyon katsayısı = 0.62), EPE ve VLT puanları (0.75), EPE ve birleşik puanlar (0.88), Cloze testi ve birleşik puanlar (0.88), VLT ve birleşik puanlar (0.93) ve Cloze testi ve VLT arasında (0.73) güçlü pozitif korelasyonlar göstermiştir. Üç test arasındaki güçlü korelasyonlar ve birleşik ölçü, bu testlerin yapısal geçerliliğinin yüksek olduğunu ve birleşik ölçünün genel yeterliliği değerlendirmede geçerli bir yaklaşım olabileceğini göstermektedir. Ayrıca, güvenilirlik analizinde yüksek olarak değerlendirilebilecek bir Cronbach's α değeri bulunmuştur (.937) (bkz Tablo 5).

Yeterlilik seviyesi, liste türü ve etkileşim terimini içeren ayrı çoklu regresyon analizleri, bu değişkenler arasındaki ilişkileri ortaya koymak için gerçekleştirilmiştir. Çalışmanın odak noktası, yeterlilik ile liste türü arasındaki etkileşimin yanlış bellek oranlarını nasıl etkilediğini incelemek olduğundan, modere edilmiş lineer regresyon analizi yapılmıştır. EPE puanları, her veri noktasından ortalama puan değerini çıkararak çoklu doğrusal bağıntıyı azaltmak için merkezileştirilmiştir. Liste türü kategorik bir değişken olduğu için, analizlerde liste türü için iki seçenekli bir dummy değişken kullanılmıştır. Regresyon denklemi şu şekildedir:

$$\text{Sahte Anı Oranları} = \beta_0 + \beta_1 (\text{Yeterlilik Düzeyi}) + \beta_2 (\text{Liste Türü}) + \beta_3 (\text{Yeterlilik Düzeyi} \times \text{Liste Türü}) + \epsilon$$

Yeterlilik düzeyinin yanlış bellek oranları üzerindeki ana etkisi anlamlı bulunmamıştır, bu da yeterlilik seviyelerinin yanlış bellek oranları üzerinde istatistiksel olarak anlamlı bir doğrudan etkisi olmadığını göstermektedir ($\beta = .0004$, $SE = .002$, $t(156) = .126$, $p = .9$), yani yeterlilik seviyelerindeki değişiklikler tek başına yanlış bellek oranlarındaki varyasyonları açıklamamaktadır.

Liste türü için katsayı anlamlı bulunmuş, liste türünün yanlış bellek oranlarını önemli ölçüde etkilediği ($\beta = .142$, $SE = .029$, $t(156) = 4.828$, $p < .001$) belirtilmiştir. Gözlemlenen pozitif ilişki, DRM listelerinin sahte anı oranlarını kategori listelerinden

daha yüksek olarak ilişkilendirdiğini göstermektedir, bu da ANOVA sonuçlarını desteklemektedir.

Önemli bir bulgu olarak, yeterlilik seviyeleri ile liste türü arasındaki etkileşim anlamlı bulunmamıştır ($\beta = .0007$, $SE = .002$, $t(156) = .308$, $p = .758$), bu da liste türünün yanlış bellek oranı üzerindeki etkisinin farklı yeterlilik seviyelerinde önemli ölçüde değişmediğini önermektedir.

Genel olarak, sonuçlar liste türünün yanlış bellek seviyeleri üzerinde belirgin bir etkisi olduğunu ancak yeterliliğin ve liste türü ile etkileşimin sahte anı oranlarını anlamlı şekilde etkilemediğini önermektedir. Modelin açıklayıcı gücü, istatistiksel olarak anlamlı olmasına rağmen, yanlış bellek üzerindeki varyansın yalnızca mütevazı bir kısmını açıklamaktadır.

Ayrı bir regresyon analizi, aynı kovaryantlarla birlikte d' seviyeleri üzerinde yapılmıştır ve tam modelin d' seviyeleri üzerindeki varyansın %13.2'sini açıkladığı bulunmuştur ($R^2 = 0.132$, Düzeltilmiş $R^2 = 0.115$, $RMSE = 599$). Liste türü d' üzerinde anlamlı bir tahmin sağlamıştır ($\beta = -0.302$, $SE = 0.95$, $t(156) = 4.051$, $p < .001$). Birleşik puan ile ölçülen yeterlilik de d' seviyeleri üzerinde anlamlı bir ana etki sağlamıştır ($\beta = .215$, $SE = .074$, $t(156) = 2.906$, 95% CI [.069, .362], $p = .004$) Yeterlilik ve liste türü arasındaki etkileşim terimi anlamlı bulunmamıştır (bkz Tablo 10). Bu bulgular, yeterlilik seviyesinin d' üzerinde anlamlı pozitif bir etkisi olduğunu, yani L2 yeterliliği arttıkça tanıma performansının da arttığını göstermektedir. Liste türünün ana etkisi, katılımcıların DRM liste türü koşulunda kategori liste türü koşulundan daha düşük performans sergilediğini önermekte, ancak ikinci dildeki yeterliliğin bu ilişkiyi anlamlı şekilde etkilemediği görülmektedir. Ayrıca, aynı kovaryantlarla birlikte karar verme kriteri © seviyeleri üzerinde bir regresyon analizi yapılmış, ancak model anlamlı bulunmamıştır. ($R^2 = 0.132$, Düzeltilmiş $R^2 = 0.115$, $RMSE = 599$).

BÖLÜM 4

TARTIŞMA

Bu çalışma, ikinci dilde yeterliliğin DRM ve kategori listeleri tarafından oluşturulan yanlış bellek oranları üzerindeki etkisini araştırmayı amaçlamıştır. Yeterlilik seviyesi yükseldikçe DRM tarafından ortaya çıkarılan yanlış bellek oranlarının, kategori

listelerinden daha yüksek olması; yeterlilik seviyesi azaldıkça ise kategori koşulunda daha yüksek yanlış bellek oranlarının görülmesi beklenmiştir.

Sonuçlar, liste türünün hedef, çeldirici ve kritik kelimeler gibi öğeler üzerindeki “evet” yanıt oranları üzerinde önemli bir etkisi olduğunu ortaya koymuştur. DRM listeleri, kritik ve çeldirici öğeler için daha yüksek “evet” yanıt oranları sağlamıştır. Liste türleri için doğru yanıt oranlarında bir fark görülmemiştir. Başka bir deyişle, DRM listeleri kategori listelerinden daha fazla yanlış alarm ürettiği görülmüştür. Bu, katılımcıların iki liste türü için karar verme kriterlerinde bir fark olduğunu gösterdiği düşündürülebilir. Tanıma performansını ölçen d' seviyeleri analizinde, iki liste türü arasında anlamlı bir farkın bulunması, bu düşüncüyü destekleyen bir bulgudur. Katılımcılar, DRM liste türünde hedef ve hedef olmayan öğeleri ayırt etmekte kategori listelerine kıyasla daha çok zorlanmışlardır. Ayrıca, c seviyeleri üzerinde yapılan regresyon analizinde liste türleri arasında anlamlı bir fark bulunmaması, liste türleri arasında gözlemlenen farkın yanıt önyargısından kaynaklanmadığına işaret etmektedir. Bu durum, DRM listelerinde kritik çeldirici kelimelerin hedef kelimelerden ayırt edilmesinin daha zor olduğu görüşünü desteklemektedir.

Mevcut çalışmada DRM ve kategori listeleri arasındaki farkın L2'de de geçerli olduğunu gösteren bulgu, literatürde genellikle L1'de DRM'in daha yüksek yanlış bellek oranlarına neden olduğu bulgusunun L2'de devam ettiğini göstermektedir. Literatürdeki mevcut bulgularla tutarlı olarak, DRM'nin oluşturduğu yanlış bellek oranı, katılımcıların ana dilinde genellikle bildirilen orandan daha düşük bulunmuştur. Ayrıca, kategori listeleri için sahte anı oranı (.43), literatürde bildirilen tipik aralıktan daha düşük bulunmamıştır, bu da L2'nin DRM ve kategori kaynaklı yanlış bellek üzerinde farklı etkiler yaratabileceğini ima etmektedir.

Yeterliliğin yanlış bellek oranları üzerindeki etkisini incelemek için, iki ayrı analiz yapılmıştır. İlk analizde, merkezi EPE sonuçları yeterlilik ölçütü olarak kullanılmış, sonuçlar, yeterliliğin yanlış bellek oranları üzerinde anlamlı bir etkisi olmadığını göstermiştir. Bu sonuç, literatürdeki bulgularla uyumsuzdur. Literatürde birçok çalışma, DRM paradigmasının katılımcının ilk veya baskın dilinde, ikinci dilden daha yüksek yanlış bellek oranları ürettiğini göstermektedir. Benzer şekilde, L2'de yüksek yeterliliğin daha düşük yeterlilikle karşılaştırıldığında daha yüksek yanlış bellek oranları ürettiği de birçok çalışma tarafından gösterilmiştir. Bu gözlemler, daha fazla deneyim nedeniyle daha güçlü kavram bağlantıları oluşması ve daha yüksek

aktivasyon seviyeleri ile açıklanabilir. Düşük yeterlilik seviyelerine sahip bireylerde, kavram ağlarının birbirine daha zayıf bağlarla bağlı olması, kritik kelimenin daha güçlü bir şekilde aktive edilmemesine neden olmakta ve bu da yanlış belleğe daha az yatkınlık oluşması şeklinde yorumlanmaktadır.

Öte yandan, hem L1 hem de L2'yi günlük yaşamda sıkça kullanan veya L2'de akademik eğitim alan katılımcılarda iki dil arasındaki yanlış bellek seviyeleri benzer olmuştur. Bu sonuç, benzer yeterlilik seviyelerinde, L2'nin sık kullanımı ve deneyiminin, yeterliliğin yanlış bellek üzerindeki beklenen etkisini azaltabileceğini ima edebilir. Çalışmadaki katılımcıların, ODTÜ'de yüksek öğrenim gören ve akademik bağlamda günlük olarak İngilizce kullanan öğrenciler olması nedeniyle dilin aktif kullanımı, yeterlilik seviyelerinin yanlış bellek seviyeleri üzerindeki potansiyel farkını etkilemiş olabileceği ve sonuçların önceki bulgulardan farklı olmasına neden olmuş olabileceği düşünülmektedir.

Liste türü ve dil yeterliliği etkileşimi üzerinde anlamlı bir etki bulunmamıştır, bu da yeterliliğin DRM ve kategori listeleri tarafından uyarılan yanlış bellek oranları üzerindeki farklı etkileri etkilemediğini gösterir. Bu sonuç, yeterliliğin sürekli bir değişken olarak değerlendirilmesinden kaynaklanıyor olabilir. Üç farklı yeterlilik ölçümü kullanılmış olmasına rağmen, EPE, VLT ve Cloze testleri gibi, katılımcıların yeterlilik seviyelerinde yeterli varyans bulunmamış ve yeterlilik seviyelerindeki farklar net bir şekilde ortaya çıkmamıştır. Literatürde yeterliliğin sahte anı oranları üzerindeki etkisini rapor eden çalışmalar, çoğunlukla öz-bildirim ölçümleri kullanmaktadır.

Araştırmanın bulguları, DRM ve kategori listeleri arasındaki farklılıkların ikinci dilde (L2) de mevcut olduğunu ve sadece anadil ile sınırlı olmadığını göstermektedir. Bu çalışma, bu farklılıkların ikinci dil bağlamında incelenen ilk çalışma olma özelliğine sahiptir. Ayrıca, dil yeterliliğinin yanlış bellek oranları üzerindeki etkisinin belirgin olmaması ve list tipi ile etkileşim etkisinin bulunmaması, yüksek ikinci dil yeterliliğinin yanlış bellek oranlarını artırdığına dair mevcut literatürle çelişmektedir. Bu tutarsızlık, yeterlilik seviyesinin henüz tespit edilmemiş bir etkisi olabileceğini önermektedir. Genel olarak, bu sonuçların, yanlış bellek mekanizmalarının anadil ve ikinci dil üzerindeki işleyişini anlamak için önemli olduğu ve bu fenomeni daha derinlemesine araştırmak için ek çalışmalara ihtiyaç duyulduğunu vurgulamaktadır.

Yukarıda belirtildiđi gibi, katılımcıların yeterlilik seviyelerindeki potansiyel deđiřkenliđi ele almak için 3 farklı deđerlendirme uygulanmıř olmasına rađmen, katılımcıların yeterlilik seviyelerinde, yanlış bellek oranlarını dođrudan etkileyecek veya liste türünün yanlış bellek oranlarını nasıl etkilediđi açısından fark oluřturmaya yetecek düzeyde varyans bulunamamıřtır. Gelecek alıřmalarda, katılımcıların daha belirgin yeterlilik farklarına sahip olmalarını sađlamak faydalı olabilir. Örneklem büyüklüğünün artırılması, mevcut etkilerin tespitini sađlamak için daha iyi istatistiksel güç sađlayabilir.

G. THESIS PERMISSION FORM / TEZ İZİN FORMU

ENSTİTÜ / INSTITUTE

- Fen Bilimleri Enstitüsü / Graduate School of Natural and Applied Sciences**
- Sosyal Bilimler Enstitüsü / Graduate School of Social Sciences**
- Uygulamalı Matematik Enstitüsü / Graduate School of Applied Mathematics**
- Enformatik Enstitüsü / Graduate School of Informatics**
- Deniz Bilimleri Enstitüsü / Graduate School of Marine Sciences**

YAZARIN / AUTHOR

Soyadı / Surname : Aykut
Adı / Name : Bahar
Bölümü / Department : Psikoloji / Psychology

TEZİN ADI / TITLE OF THE THESIS (İngilizce / English): Effects of Proficiency in a Second Language on False Memories Produced by Associative and Categorical Procedures

TEZİN TÜRÜ / DEGREE: Yüksek Lisans / Master Doktora / PhD

1. **Tezin tamamı dünya çapında erişime açılacaktır. / Release the entire work immediately for access worldwide.**
2. **Tez iki yıl süreyle erişime kapalı olacaktır. / Secure the entire work for patent and/or proprietary purposes for a period of two years. ***
3. **Tez altı ay süreyle erişime kapalı olacaktır. / Secure the entire work for period of six months. ***

** Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. / A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

Yazarın imzası / Signature **Tarih / Date**

*(Kütüphaneye teslim ettiğiniz tarih. Elle doldurulacaktır.)
(Library submission date. Please fill out by hand.)*

Tezin son sayfasıdır. / This is the last page of the thesis/dissertation.