Are children mindful decision makers?

Investigating underlying mechanisms of Turkish-learning children's recency bias

By

Zeynep Adıgüzel

A thesis submitted to

the Graduate School of Social Sciences and Humanities in Partial Fulfillment of

the Requirements for the Degree of

Master of Arts

in

Psychology

September 2021

DEDICATION

To all children of the world and to make their voice louder

STATEMENT OF AUTHORSHIP

This thesis contains no material which has been accepted for any award or any other degree or diploma in any university or other institution. It is affirmed by the candidate that, to the best of her knowledge, the thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signed Zeynep Adıgüzel

ABSTRACT

Are children mindful decision makers?

Investigating underlying mechanisms of Turkish-learning children's recency bias

This study examined (a) whether Turkish-reared preschoolers (n = 59, Mage = 53.37 months, SD = 10.11) showed recency bias in a flexible word order language, and (b) how their recency bias was related to individual differences in working memory and verbal skills. Children were presented with a decision-making task in which they responded to questions about an imaginary character's decisions. They listened to questions that had both choice options at the beginning, at the end or first option being at the beginning while the second at the end of the sentence. Then, children were given tests to assess their working memory (WM) and verbal skills. We found that children exhibited recency bias, and it decreased as their age increased. Children's WM skills, not language skills, were a strong predictor of recency bias, even after controlling for age. Whether the options were familiar or located at different places within the question did not have an effect alone. However, slope analyses showed that their interactions with each other and children's WM skills may have yielded a significant decrease in recency bias. That is, as children aged, their recency bias decreased when responding to unfamiliar options if both were encountered at the beginning of the question. Recency bias decreased as their WM skills improved when both options were encountered at the beginning of the questions. This study investigated recency bias systematically in a free word order language with respect to individual differences, suggesting that children are not always mindful decision-makers. These findings have implications for unveiling children's response biases and ways to elicit effective communication with them.

Keywords: recency bias, working memory, response bias, decision making, Turkish word order

ÖZETÇE

Çocuklar Bilinçli Karar Vericiler Midir? Türkçe Konuşan Çocuklarda Sonralık Eğilimini Altında Yatan Faktörlerin Araştırılması

Bu araştırma, (a) serbest sözdizimine sahip bir dil olan Türkçe'yi öğrenen çocukların (N = 59, M = 53.37, SS = 10.11) sonralık eğilimi (*recency bias*) gösterip göstermediğini ve (b) bu eğilimin çocukların işleyen bellek ve dil becerilerindeki bireysel farklılıklara bağlı olup olmadığını incelemiştir. Çocuklara kurgusal bir karakterin kararları hakkında iki seçenekli soruların sorulduğu bir test verilmiştir. Çocuklara yöneltilen bu soruların üçte birinde seçeneklerin ikisi de cümle başında, üçte birinde iki seçenek de cümle sonunda, kalan üçte birinde de bir seçenek cümle başında iken diğeri en sonda yer almaktadır. Seçeneklerdeki objelerin yarısı bildikleri kelimeler yarısı da daha önce hiç duymadıkları yeni oluşturulan kelimelerden seçilmiştir. Bu karar verme testinden sonra çocuklara işleyen bellek ve dil becerilerini ölçen testler verilmiştir. Sonuçlar, çocukların sonralık eğilimine sahip olduğunu ve bunun yaş arttıkça azaldığını gösterdi. Çocukların sonralık eğilimlerinin, işleyen bellek performansı tarafından, yaşın da ötesinde, yordandığı, ancak dil becerilerinin buna bir etkisinin olmadığı bulunmuştur. Seçeneklerdeki objelerin bilinip bilinmediği ya da seçeneklerin cümlenin farklı yerlerine konulması tek başlarına bir etkide bulunmazken; birbirleriyle etkileşimi anlamlı bir etki gösterdi. Bu da çocukların yaşları arttıkça ve seçeneklerin ikisi de başta olunca ve bilmedikleri seçenekler hakkında sorular cevaplarken sonralık eğiliminde azalma görüldü. Bireysel farklılıklar açısından da çocukların işleyen bellek performansları arttıkça ve seçeneklerin ikisi de başta olunca sonralık eğiliminde azalma görüldü. Bu araştırmada, serbest sözdizimine sahip dili konuşan çocukların bireysel farklılıklara göre sistematik olarak sonralık eğilimi incelenmiştir. Böylelikle çocukların her zaman bilinçli karar verici olmadıklarını önermek mümkündür. Bu bulgular, çocukların karar verme süreçlerinde yaşadıkları eğilimleri ortaya çıkarma ve önleme açısından önemlidir.

Anahtar kelimeler: sonralık eğilimi, işleyen bellek, yanıt yanlılığı, karar verme, Türkçe serbest sözdizimi

ACKNOWLEDGEMENTS

2020-2021, what a time to write a thesis under the challenging times of the pandemic!

These years do not need an introduction from me but I have come a long way.

I would like to thank some of the people without whom I would not have been able to complete this research and my degree!

I would first like to thank my brilliant thesis supervisor, Tilbe Göksun. She has always been a mentor for me, even before I was her graduate student. I have witnessed her dedication to her work, fairness and affection towards every student and admired her genuine heart. Throughout this thesis project, there were times that things seemed too difficult or even impossible to me (try collecting online data with children!). However, she has always guided me wisely and empathetically, and encouraged me to try harder. She also taught me to "take a break" to come back stronger, which made me a better researcher and a happier person. She is a true role model for me, in whatever work I do.

I want to thank Sami Gülgöz, the other person who had the most impact on me. I have been a member of the KURAM Lab since my freshman year at Koç University. He was always there, not only for me but for each of his students. He supported me greatly to build on the conceptual issues of this thesis and helped me to look from a more creative perspective. He made me feel truly confident in my capabilities and the things that I can do. I am forever grateful for this.

I would also like to express my gratitude to Aslı Aktan-Erciyes, for her continuous support and guidance. She was guiding my research even before this project turned into a thesis. She has provided me with critical feedback and helped me to ground my work stronger theoretically. I take pride in having been trained under the guidance of these world-class scientists.

Furthermore, many thanks to Demet Özer, who gave me invaluable support in statistics. Her insight and knowledge helped me enormously throughout this research. I also want to thank the undergraduate research assistants Gülfem Dikici, Nurgül Arslan and Mine Yürekli for their enormous help and hard work. I am very glad that I had the chance to work with you.

I am also grateful for the lab environments that enabled me to grow as a researcher. I want to thank the KURAM and Language and Cognition lab members, my dearest friends and colleagues, for creating an enjoyable work environment and solidarity of friendship.

I want to express my biggest thanks and gratitude to my beloved family. They have always stood by my side with such unconditional love. My father Şaban and mother Nurten, I feel enormously lucky to be your little daughter. I cannot thank them enough for giving me the opportunity to pursue every dream I had. Haluşum, I am grateful that you are in my life. Thank you for supporting me in every step of my life.

Kadir, I could not have done this without you. My amazing boyfriend and partner in crime whom I admire for the brilliancy of his mind and kindness of his heart. You have listened to me talking about my thesis so many times that I am certain you could respond to all the questions I had during the defence. But probably I would have complained to you later for not answering questions as a psychologist but as a mathematician or an engineer. You saw all my frustration and tears. Thank you for your patience and encouragement, and for even putting up with me being grumpy in times of despair. I could have not asked for more.

My friends. My other family. My neighbours at MEDENİYET! You all know what it took. From the bottom of my heart… Thank you Çilli and Koko, for just existing, really. Thank you Nuriş for being the best mom to Çilli and Koko, and for hosting us even when you are not home, Yasemin for countless work-up-until-nights and coffee parties, Yavuz for day-time-study-buddies and tantuni parties, Burak for your expertise in technical challenges I had.

Murti, Duygu, Ata, Aykut, Arda and Deniz for being there for me in any way that they could. I wish I could name everyone.

Thank you Ezgi for being “that friend in academia that everyone should have”! We have become the pushing force for each other in the hardest times, especially in our Zoom study dates. Thank you Dilan and Çiğdem for sacrificing your time to listen to my defense talk more than once! Thank you Ceylan for bringing your design talent into my research. Thank you Işıl for giving your voice to my weird but cute imaginary characters.

Last but not least, I would like to express my sincere gratitude to the whole Koç University community. I have spent 8 years of my life in this beautiful university. Stepping my foot into this campus has changed my life. In this institution, I have grown in knowledge and experience as an individual, researcher and psychologist. Long story short, I have been a proud member of Koç University and always will be.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

To learn what children think or know, adults ask them various questions in both daily and formal situations (Okanda et al., 2012). For example, asking children questions like “Did he have a bicycle or motorcycle?” or “Do you prefer apples or bananas?” limits them to choose from one of the given alternatives (Behzadnia & Mehrani 2018). During speech comprehension, children need to understand and retain it until the whole message is received (Cowan, 2005). This process is related with children’s working memory capacity (Just et al., 1996; Just & Carpenter, 1992; Mehrani & Peterson, 2017; Sumner et al., 2019) and language skills (Behzadnia & Mehrani, 2018; Mehrani & Peterson, 2017; Sumner et al., 2015). Hence, the success of comprehension may depend on the interplay between the demands of language processing and what children’s cognitive systems can afford.

When linguistic processing demands exceed processing capacity, children tend to exhibit certain response tendencies (Fritzley et al., 2013; Moriguchi, Okanda & Itakura, 2008) For example, *the recency bias* refers to preferring the most recently mentioned choice out of the two presented ones (Mehrani & Peterson, 2015). While some researchers found no bias for preschoolers, others showed a mild or strong bias for younger children. Their bias increases when asked about unfamiliar objects, which is suggested to be underlined by the limited working memory of the youngsters (Sumner et al., 2019). Another factor associated with the recency bias can be language-specific properties such as word order. Turkish with its flexible word order provides good sentential context to test whether syntactic variations in a language interact with children’s limited cognitive resources and hence affect their recency tendencies.

The present study investigates whether and how 3- to 5-year-old children have recency bias in their responses to two-alternative forced-choice questions in Turkish, a flexible word order language. With this aim, we administer a decision-making task in which children will make decisions about an imaginary character’s preferences. We also measure their working memory and verbal skills. We ask two main questions about children’s recency bias and

potentially related factors. First, is 3- to 5-year-old Turkish-learning children's recency bias influenced by the flexible word order of Turkish? Second, does it depend on individual differences in their working memory and verbal skills?

1.1 Children's Recency Bias

When asked to choose from one of the two alternatives, children's tendency to choose the most recently mentioned choice is called *recency bias*. These two-alternative forced-choice questions are used in applied settings such as interviews done by child protective services, for example, on child sexual abuse (Davies et al., 2000) and experimental paradigms examining social and cognitive development (Fritzley, 2000). Recent evidence suggests that children's responses might be open to biases and might not be reliable (Behzadnia & Mehrani, 2017, 2018; Brady et al., 1999; Ceci & Burak, 2006, Fritzley & Lee, 2003; Fritzley et al., 2004; Okanda et al., 2013; Sumner et al., 2015, 2019). Children's tendency to reply yes (i.e., yes-no bias) has been the most investigated bias while yielding inconsistent results across studies (Okanda & Itakura, 2010; Fritzley & Lee, 2003; Fritzley et al., 2013; Behzadnia & Mehrani, 2018; Peterson & Grant, 2001), On the other hand, very few studies investigated recency bias in forced-choice questions (Sumner et al., 2015; Mehrani & Peterson, 2015), which is the focus of the present paper.

Earlier evidence suggests the existence of recency bias in children. For example, Mehrani and Peterson (2015) showed 3- to 5-year-old Persian-speaking children a video about animals and asked two-alternative forced-choice questions about it. Children exhibited a recency tendency that weakened as children aged. When the questions were asked in the opposite order (*"Was it spring or winter?"* was asked as *"Was it winter or spring?*"), similar results were obtained (Mehrani & Peterson, 2015). A follow-up study compared Persian-, Kurdish- or English-speaking children who were between the ages of 2 and 5. They questioned children about objects and manipulated children's familiarity with objects (familiar vs.

unfamiliar objects) (Mehrani & Peterson, 2017). Regardless of their age and language background, children demonstrated a recency bias. Additionally, children's biases were stronger when asked about unfamiliar objects and became weaker as they aged (Mehrani & Peterson, 2017). Behzadnia and Mehrani (2018) applied a very similar paradigm with Persian-Turkish bilingual 5- and 6-year-old, half of the questions were asked in Persian and half were in Turkish[1] (using Persian alphabet and a different dialect than Turkish used in Turkey). Their results replicated previous studies in addition to showing no effect of language in which questions were asked. Recently, Sumner and colleagues (2019) presented English-learning toddlers with a decision-making game, involving forced-choice questions about imaginary characters. The findings again replicated earlier studies, indicating a strong recency bias that became weaker as children's age increased. However, object familiarity in this study did not influence children's responses.

In sum, younger children have a strong recency bias both for familiar and unfamiliar choices whereas older children have the biases only when the choices are unfamiliar (Mehrani & Peterson, 2017; Behzadnia & Mehrani, 2018). Additionally, 3 years of age might be a developmental transition period for response biases (Fritzley & Lee, 2003; Mehrani & Peterson, 2015; Steffensen, 1978) that can be universal across languages (Behzadnia & Mehrani, 2018; Mehrani & Peterson, 2017). However, previous studies overlooked language-specific factors that can potentially influence children's recency tendencies.

Word order is one important and meaningful cue in sentence comprehension since it portrays information about events and their participants concerning those events. Previously studied languages had different word orders, for example, SOV in Persian (Mehrani & Peterson, 2015), SVO in English (Sumner et al., 2019). Yet, in the experimental conditions,

[1] Behzadnia and Mehrani (2018) notes that these languages share several commonalities such as lexicon due to their speakers being graphically close and in constant interaction. However, they reasoned that other linguistic differences such as intonation patterns, word order, and other syntactic differences could have affected children's biases.

they used only one order. Sumner et al. (2019) varied question types in terms of where the choices were positioned but did not find an effect of question type on children's recency tendencies. However, their question stimuli were incomparable across conditions due to fixed word order in English. Turkish, on the other hand, has a flexible word order with 6 different variations of S, O, and V, though SOV is the canonical one. Although these different word orders have pragmatic differences, Turkish speakers can position options in different places in the sentence while preserving the meaning. All these aspects of Turkish may not only affect the way children demonstrate recency tendency but also interact with cognitive individual differences.

1.2 Individual differences in children's cognitive resources

Beyond language-related factors, little is known about whether and how recency bias is associated with individual differences in children's cognitive skills. One related system is working memory that refers to the brief and limited capacity to store and manipulate information (Baddeley & Hitch, 1974; Just & Carpenter, 1992), for not only final but also intermediate parts of a comprehension process (Just & Carpenter, 1992).

Most of the working memory research was conducted following the tripartite model of Baddeley (Alloway et al., 2004; Baddeley & Hitch, 1974). In this model, the first sub-system is the phonological loop, which was specialized in auditory information. It serves both in holding its memory trace and preventing it from decay by rehearsing (Baddeley, 2007). The other subsystem is specialized in visuospatial information and performs similarly. Episodic buffer was later added to the model in need of including the influence of long-term memory. However, the episodic buffer has not been investigated as much as other WM sub-systems, especially in developmental psychology (Boyle et al., 2013) They are both limited in capacity and "slave systems" of the domain-general central executive system. Although having a limited capacity as the other sub-systems, the central executive directs the distribution of finite WM

resources and controls the attentional systems (Baddeley, 2007) by three main functions: inhibiting more prominent but less relevant information (inhibition), shifting across different tasks and mental sets (mental set shifting), and updating and monitoring information (updating) (Miyake et al., 2000). Therefore, its performance is generally measured with tasks that involve both information retention and manipulation, which are referred to as complex or executive-loaded WM tasks, as opposed to storage-only tasks that only require retention but not manipulation of information (Blom et al., 2014).

WM system undergoes a protracted development. WM improves 4 years onward (Gathercole et al., 2004), which is related to structural and functional maturation in frontal-parietal association cortices (Anderson et al., 2001). As children age, more activation is observed in these regions (Kwon et al., 2002) that matures throughout adolescence into young adulthood (Conklin et al., 2007; Luciana et al., 2005). Hence, due to WM-based limitations, young children may not process some early word representations to integrate and relate to later ones quickly enough (Just & Carpenter, 1992), which would jeopardize the success of understanding the questions and answering appropriately (Cowan et al., 2005).

Although some studies examining children's recency bias suggested that WM skills might be an underlying mechanism (Mehrani & Peterson, 2017; Sumner et al., 2019), surprisingly only one study tested its influence on children's responses. They found that WM did not predict children's tendency to choose second response choice. However, the bias was stronger when the response choices were longer and was limited to children's verbal responses, which led them to conclude that WM limitations could be underlying mechanisms for children's response biases (Sumner et al., 2019). It is also possible that their working memory measure (Shriberg et al., 2009) was too difficult for 3- to 4-year-old children since it only consisted of non-words such as dama, dabama, danabama. WM performance of young children is mostly measured by span tasks that are also related to children's sentence comprehension in

reading and listening tasks (Booth et al., 2000), even when the effect of vocabulary knowledge was ruled out (Swanson, 1992). Based on the empirical evidence and the given functions of WM, we expect WM resources to play a central role in the recency tendencies of children.

Children's verbal or language skills also improve drastically during preschool ages. Vocabulary knowledge is a good predictor of speech comprehension (Jun Zhang & Bin Anual, 2008; Malatesha Joshi, 2005; Martin-Chang & Gould, 2008; Moghadam et al., 2012), both concurrently and longitudinally (Nation & Snowling, 2004). Children's immature verbal skills associate with their yes-no biases (e.g., Moriguchi et al., 2008; Okanda & Itakura, 2010); however, only a few studies addressed recency bias in relation to verbal skills. While Mehrani and Peterson (2017) discussed that recency bias might be due to immaturity in language development in areas like sound system, syntax, semantics, and pragmatics, etc., Sumner et al. (2019) found no association with children's verbal skills. In the process of language development, children can utter a word without a complete lexical representation of it (Mehrani & Peterson, 2017; Sumner et al., 2015). For example, children as young as age 2 may not be able to articulate what they know even when they have the knowledge (Berthier et al., 2000), which might influence their recency tendencies. The present study will also investigate whether children's recency bias is associated with their verbal skills.

In sum, children exhibit a recency bias that is associated with individual differences in cognitive resources. WM system has an active role in maintaining and processing linguistic information as the sentence unfolds its meaning (Borovsky et al., 2012; Emmorey et al., 2017). However, children's ability to keep items in mind over short delays matures late in development (Alloway et al. 2006). Their vocabulary also undertook significant changes that may predict children's response tendencies in addition to their WM skills. Thus, children with lower WM functioning and vocabulary size, especially younger ones, may have difficulty in

keeping the whole message in mind, which would underlie the tendency to choose the last heard choice.

1.3 The present study

The present study investigates 3- to 5-year-old children's recency biases in a language with a flexible word order and its associations with children's WM and verbal skills. Children were given a decision-making task in which they made choices regarding preferences of imaginary characters to investigate their recency biases. Our study contributes to previous research in three ways. First, earlier studies either had choices at the end of the sentence in all trials (Mehrani & Peterson, 2015) or incomparable variations in terms of the position of choices (Sumner et al., 2019). Since Turkish allows different word orders, it provides a particularly interesting case to test whether variations in word order will affect recency bias. We formulated our questions by changing the position of options while keeping the rest of the syntactic structures the same across conditions. We generated three different and comparable question types: 1) Options at the end *"Hangisini saklamayı tercih eder Aymi elmayı mı üzümü mü*? ("Which one Aymi would prefer to hide, *apple or grapes*?") 2) One option at the end and one at the beginning *"Elmayı mı saklamayı tercih eder Aymi üzümü mü?* ("*Apple,* which one Aymi would prefer to hide, or *grapes?")* 3) Options at the beginning *"Elmayı mı üzümü mü saklamayı tercih eder Aymi?"* ("*Apple or grapes,* which one Aymi would prefer to hide?")[2]. Second, as another related question, previous studies presented children with familiar and unfamiliar objects as choices. However, unfamiliar objects like "tweed, "ganache," "teepee" in Sumner et al. (2019) and "plastic," "laundry catcher," "paper filter for green tea" in Okanda et al. (2012) may not be truly novel to children. Hence, to test whether children's biases depend on the familiarity of options, we used novel objects created for this study as unfamiliar choices. Third, previous studies documenting response biases have not fully tested individual

[2] The literal translations of the Turkish sentences are in parentheses

differences in cognitive skills of children (Moriguchi et al., 2008). The present study aims to provide a more complete developmental picture of the phenomenon.

First, previous studies have found that children around the ages of 3 and 5 showed a recency tendency (Behzadnia & Mehrani, 2019; Mehrani & Petereson, 2015; 2017). We aimed to replicate this consistent finding. Additionally, since some argued that recency bias might be universal across languages (Behzadnia & Mehrani, 2018; Mehrani & Peterson, 2017), although tested in just a few languages, our goal is to show children's recency bias in a different language. In line with previous studies (Mehrani & Peterson, 2015; 2017, Sumner et al., 2019), we expected that, overall, recency bias will be weaker in older children. For example, Mehrani and Peterson (2005)'s results yielded strong recency bias while 3-year-olds scored higher among 3- to 5-year-olds. Upon this consistent finding, evidence suggests that children may undergo a developmental maturation in terms of answering two-alternative forced questions appropriately (Fritzley et al., 2009; 2013).

Second, we investigate possible language-specific factors that are related to question properties. Regardless of age, we also expect recency bias will be the strongest when the distance between 2 options is largest (beginning-end condition). We do not predict a difference in responses between the other two conditions (beginning-beginning and end-end conditions). Sumner et al. (2019) manipulated word order and found no effect since English may not an appropriate language to test word alternations. Thus, we expect to find such an effect in a free word language. In the beginning-end condition, having the largest distance between two options may take a greater effort to retrieve the earliest one. Therefore, we predict that children will baffle at most when comprehending this type of question, leading them to experience a more pronounced recency bias.

Furthermore, although recency bias fades away as children age, previous studies indicated that younger children had recency bias for both familiar and unfamiliar options;

however, older children had only for unfamiliar options or no bias at all (Behzadnia & Petreson, 2017, Mehrani & Peterson, 2017). Although 2-year-olds in Sumner et al. (2019)'s study were not affected by the familiarity of options, other evidence points to the otherwise, which may suggest that retrieving the earliest option was more difficult especially since the child did not have prior knowledge about it. Thus, we expect that as children's age increase, there will be a decrease in recency bias when children responding to familiar options compared to unfamiliar ones. We also expect this association to be stronger when a child waited the longest to hear the second option after hearing the first one (beginning-end condition) (a 3-way interaction between age, familiarity, and question type).

Third, we investigate the role of individual differences in the cognitive skills of children as possible underlying mechanisms of recency bias. In line with previous studies (Mehrani & Peterson, 2017; Sumner et al., 2019), we expect that children's WM skills will play a central role on their recency bias. Given that WM functioning is key to hold and maintain earlier information and then integrate and relate with later ones (Cowan et al., 2005, Just & Carpenter, 1992), children who have higher WM performance should be less prone to show recency bias.

Moreover, although Sumner et al. (2019) found that verbal skills did not relate to recency bias, which might be due to using an adult-level WM task for young children, some studies discuss verbal skills to have an underlying influence on recency bias (Moriguchi et al., 2008; Okanda & Itakura, 2010). Additionally, children who have poorer language skills are more likely to be suggestible since they may be more likely to respond to a question that they do not fully comprehend (Ceci & Bruck, 2006). Thus, we expect that children with lower verbal scores will exhibit more pronounced recency bias.

Last, we expect that children's improving WM skills will help them overcome the effect of word order. That is, the effect of word order will be weaker as children's WM performance increases. Our earlier prediction, experiencing the largest distance between two options leading

to a stronger recency bias (beginning-end condition), will be weaker as children's WM performance increases.

CHAPTER 2: METHOD

2.1 Participants

We advertised our study via social media channels of the lab, and posts on parenting channels on Instagram and Facebook. We recruited 70 Turkish-learning children. Among these children, we excluded three 2-year-olds as they failed to communicate with the experimenter and could not proceed with the task. Six children were also excluded because they failed to follow instructions and tasks terminated without success. Two more children were excluded due to side bias (i.e., always choosing the object presented at one side of the screen). Thus, our final sample included 59 Turkish-speaking children (33 girls) aged between 35 and 71 months (*M*age = 53.37 months, *SD* = 10.11). Informed consent was obtained from the parents of the participants via Zoom. Koç University's Institutional Review Board approved the study (Project name: Children's Recency Bias to Forced-Choice Questions, Protocol No: 2020.004.IRB3.003)

We determined the sample size by using G*Power software. It was calculated by defining effect size as .25, the alpha error probability of .05, and power as .95. All participants were native Turkish learners. All sessions were conducted online via Zoom.

2. 2 Measures

2. 2. 1 Decision-Making Task

We developed a decision-making task to investigate whether children show any recency bias when answering two-alternative choice questions, inspired and adapted from Mehrani and Peterson (2015) and Sumner et al. (2019). In this task, children were provided with imaginary characters (named either Aymi or Tibi) and asked to choose what these characters would prefer among the two presented options.

The stimuli consisted of two sets of questions. Each set had randomly presented 24 questions with two choices (see Appendix A for and the entire list of choices and question stimuli). For each question, there were 2 phases: The presentation phase and the test phase. The

presentation phase included each image choice shown at the center of the screen and accompanied by an audio recording of the choice. The test phase included the character image at the upper-middle of the screen and two choices at the bottom-corners of the screen, accompanied by the voice recording of the question (see Figure 1 for a sample trial sequence). We used the pre-recorded voice of a female narrator. During recording, we made sure to use similar intonations in narrating questions while keeping the tone natural as in face-to-face communication.

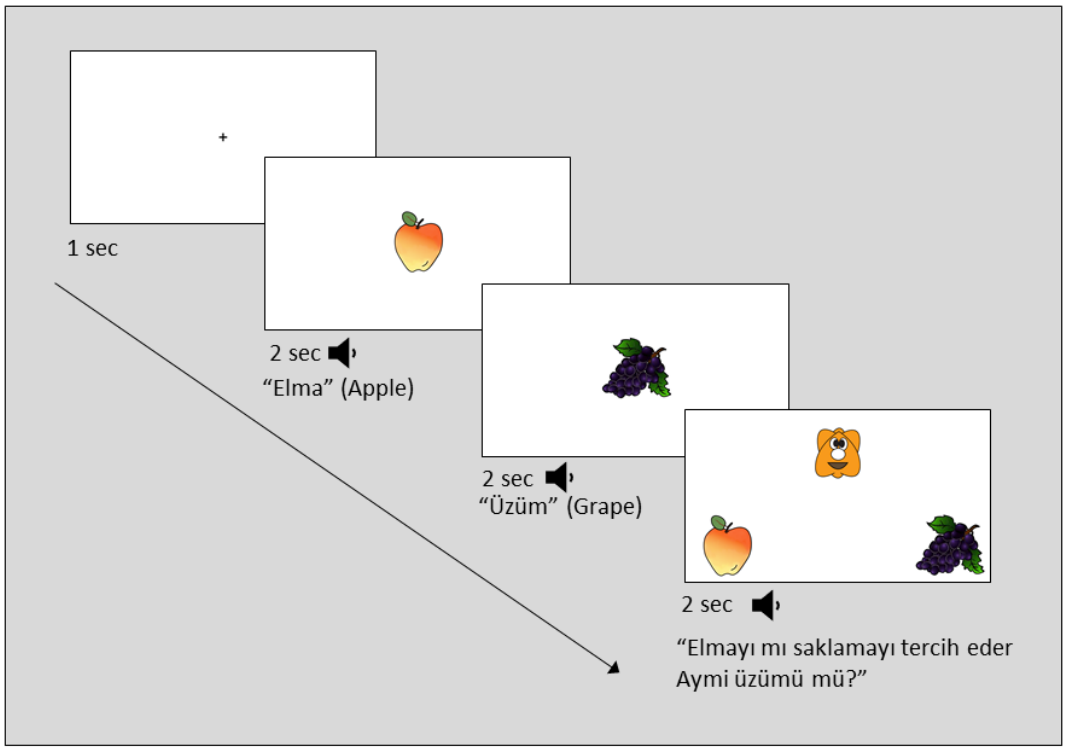

Figure 1. A sample trial sequence and timing of phases within a trial. The presentation phase included each image choice (apple and grape in a sequence). The test phase included the character and both choices, paired with a single auditory question.

In both sets, questions varied in the familiarity of the objects in choices and the position of these choices within the sentence (see Figure 2 for sample stimuli). About familiarity, half of the questions involved choices known by 2-year-olds, thereafter familiar choices, (Göz et al., 2007; Acarlar et al., 2008) (e.g., üzüm, elma, karınca, televizyon, otobüs) and the other involved questions with unfamiliar/novel choices (e.g., terka, maka, ilmasan,

pızadeni, bikirum) thereafter unfamiliar choices. To construct novel words for the unfamiliar words, we first prepared a pool of commonly used Turkish syllables and then randomly selected syllables from that pool. Importantly, the unfamiliar choices matched with the familiar ones in terms of syllable number (see Appendix A for the entire familiar and unfamiliar choices). After selecting words for familiar and unfamiliar choices, we constructed their visual images from the picture database for Turkish-speaking children prepared by Raman et al. (2014). For familiar choices, we directly used the corresponding images in the database. However, for unfamiliar choices, we followed a novel method. Other pictures in the database, that were different than what we used for familiar choices, were reconfigured by a designer. We reconfigured the pictures while sticking to the basic Gestalt principles as in we maintained the figure-ground ratios from the original picture until they did not resemble any known object. To select among these novel set of pictures, we conducted a pilot study. We asked adult participants to indicate what they saw in these pictures. They responded either "I do not know" or wrote what they thought it resembled. Then, they were asked to rate the similarity of this novel image to the one they thought it resembled on a 1 to 5 Likert scale (1= Not at all similar, 5=Extremely similar). Among 32 reconfigured images, we first eliminated those that had a mean rating of above 3. We then eliminated those that were thought to resemble similar objects by different participants because it signalled that those objects looked like something else that we did not anticipate but participants agreed on. This way, we selected images picturing novel objects or objects that were thought to be the least similar to something familiar/known (see Appendix B for the images for familiar and unfamiliar choices).

In terms of choice positions, one third of the questions had both choices at the beginning (e.g., *Elmayı mı üzümü mü saklamayı tercih eder Aymi?"* ("Apple or grapes, which one Aymi would prefer to hide?"), one third had both choices at the end (e.g., *"Hangisini*

saklamayı tercih eder Aymi elmayı mı üzümü mü? ("Which one Aymi would prefer to hide, apple or grapes?"), and the last one third had the first choice at the beginning and the second choice at the end (e.g., *"Elmayı mı saklamayı tercih eder Aymi üzümü mü?* ("Apple, which one Aymi would prefer to hide, or grapes?"). Two sets of questions only differ in the order of choices presented and the character whom the questions are about. For example, the question in one set *"Elmayı mı saklamayı tercih eder Aymi üzümü mü?* ("Apple, which one Aymi would prefer to hide, or grapes?") is asked as "*Üzümü mü saklamayı tercih eder Tibi, elmayı mı?"* (Grapes, which one Tibi would prefer to hide, or apples) in the other set.

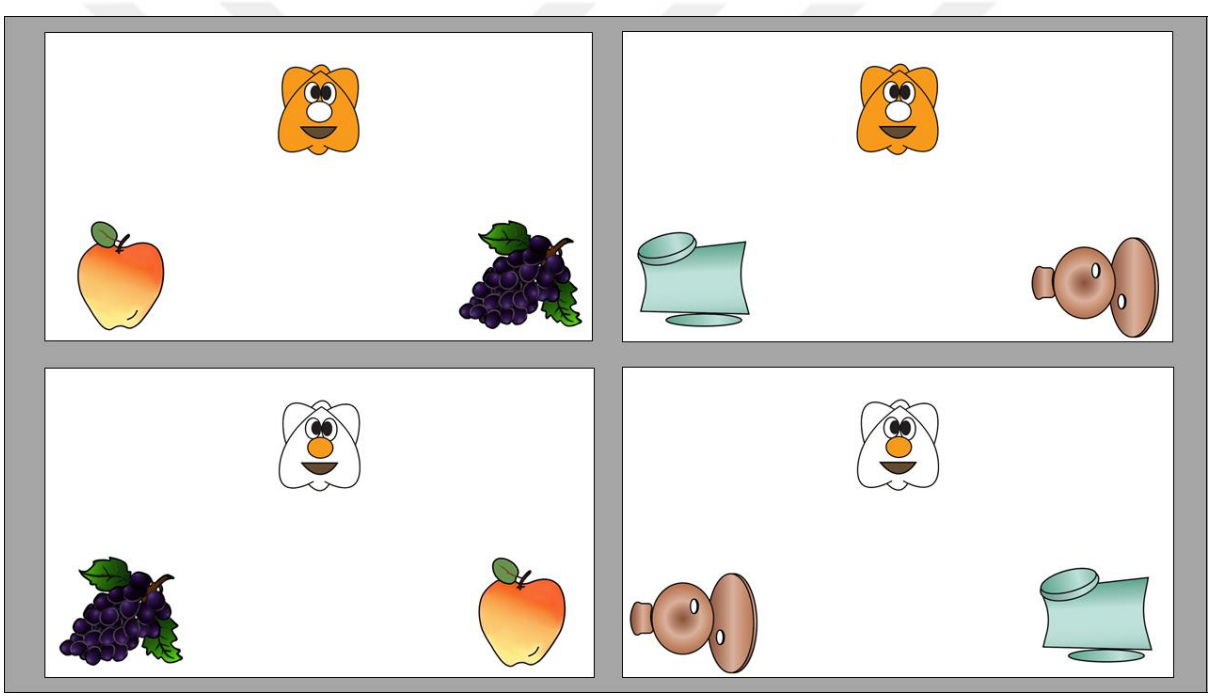

Figure 2. Sample stimuli. Questions with familiar (right panel) and unfamiliar objects (left panel).

There were 3 practice trials where the experimenter made sure that the child understood the task (see Figure 3 from the child's perspective during the experiment). The practice trials served as calibration trials in the sense that children's responses were deducible from his or her hand and arm movements. After the practice session, the experimenter introduced a character (either Aymi or Tibi) and asked the child if she could help this character's choices. After asking the first set of questions about one character, the experiment moved to the second set of

questions about the other character. The order of question sets (Aymi or Tibi questions) was counterbalanced across participants. For the decision-making task, the number of second choices was summed and its proportion was calculated with respect to the total number of questions answered. This percentage score was named as the *recency score.*

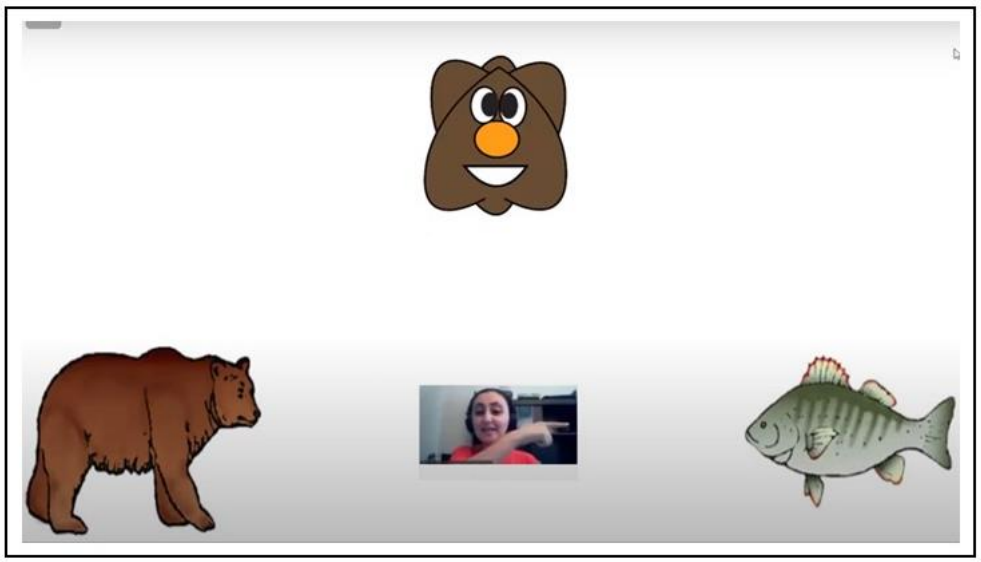

Figure 3. Child's viewpoint during the practice trial

2. 2. 1. 1 Coding

We coded participants' responses in the decision-making task. As revealed in the pilot testing, children often used both verbal and nonverbal (e.g., pointing) responses. We first coded whether responses were verbal or nonverbal. Then, we assessed the relations between these two types of responses: (1) matched vs. mismatched (i.e., referring to the same choice), (2) temporally synchronous vs temporally asynchronous (i.e., their temporal synchrony). With this coding schema, we could compare responses across domains. We coded children's responses for each trial based on this coding schema. For example, a child can give only a verbal response or non-verbal response. It is also possible that a child can give both verbal and non-verbal responses for a single trial that can refer to the same choice (matched and temporally synchronous) or to different choices (mismatched and temporally synchronous). Verbal and nonverbal responses can also be made consecutively while referring to the same choice

(matched and temporally asynchronous) or to different choices (mismatched and temporally asynchronous). In addition, we excluded trials in which a child gave a response before the whole question was being uttered. We have also excluded trials in which it took too long for the child to respond (e.g., if looked around and waited for a while).

The first author and a research assistant coded 20% of the data independently for reliability analysis (Cohen's Kappa = .89). Once disagreements were solved via discussion, they shared coding the rest of data.

2. 2. 2 Digit Span Tasks

We measured children's phonological/verbal working memory with the subset of the Turkish adaptation of WISC-4 that includes two-digit span tasks: forward and backward digit span tasks. Turkish standardization, validity, and reliability of these tests have been established (Savaşır & Şahin, 1995). In the forward digit span task, children were first provided with random digits and later were asked to repeat the exact digit array. The length of the digit array was increased by one digit every two trials. After two practice trials, children were given one point for every time they accurately reproduced the array. In the backward digit span task, children were asked to repeat the given array in the reverse order ("5, 3, 8" would become "8, 3, 5"). After two practice trials, children were given one point for that reversed trial. In both digit span tasks, the length of the digit array was increased by one digit every two trials. The maximum score for both span tasks was 16. The tests ended when children could not recall both sequences of a given array length.

2. 2. 3 Word Span Tasks

We used a second set of working memory task that was developed for this present study based on the same principles of those previously constructed (e.g., Akoğlu & Acarlar, 2014; Blake et al., 1994); Pickering & Gathercole, 2001) with two additions. In this task, both the word length (number of syllables) and set size (number of words per sequence) were

manipulated. At each level, the number of syllables and words increased by one. In previous studies, the word span tasks used nonsense words that are one syllable long (Brener, 1940, added) and one to four syllables (Akoğlu & Acarlar, 2014) or to five (Santos et al., 2006). Meaningful words that are one syllable long (Ünal, 2018; Engle, Cantor & Carullo, 1992) and one to two syllables (Engle et al., 1991) such as animals (Blake et al., 1994) were also used. Set size usually increases and is typically presented in ascending order (of set size). Also, there are multiple trials for each set size (May et al., 1999). Overall, previous span tasks had either an increasing set size over trials (Saito & Miyake, 2004; Pickering & Gathercole, 2001; Blake et al., 1994) or increasing word length across trials but keeping the number of words per trial constant (Archibald & Gathercole, 2006; Gathercole & Pickering, 2000; Gathercole & Baddeley, 1996) (e.g., sets of 10, each containing two, three, four and five syllables- Gathercole & Baddeley, 1996). Incorporation of the two is rare. For example, Hulme et al. (1991) constructed a pool of 2, 3, and 4 syllables long words and created word sequences from 2 to 8 items by randomly choosing words. Additionally, Barrouillet et al. (1995) sampled animal names to create an increasingly long series of animal lists. However, in both, the word length per set does not increase as the size does. With our WM measure, children's WM skills can be measured more comprehensively and accurately since it not only represents its functioning as of an increasing number of items (words) but also syllables of items (words).

This WM task included forward and backward word span. In the forward word span, children were asked to recall a given set of words in the presented order. In the backward word span, children were asked to recall the given sequence of words in the opposite order. In both tasks, there were 2 different sequences of words in each level of sequence size. The smallest sequence size is 1 and the longest one is 4. At each level, the number of syllables and items increases by one. An example of the forward word span is the following (see Appendix C for the entire stimuli list):

1. Fil (Elephant)
2. Yangın – Şeker (Fire-Sugar)
3. Süpürge- Şemsiye- Telefon (Broom- Umbrella-Telephone)

After two practice trials, children were given one point for every time they accurately reproduced the array in a forward or backward fashion. The test was terminated when a participant inaccurately recalled both sequences of a given set. The maximum score for both span tasks was 8.

2. 2. 4 Receptive Vocabulary Task

To measure children's receptive verbal skills, we used the Turkish Receptive Language Test (TIFALDI-R) (Berument & Güven, 2010). This test is used with 2- to 12-year-old children and contains 104 items which include 4 pictures each. In our online adaptation, we prepared a PowerPoint presentation and the experimenter screen-shared it during the testing via Zoom. First, once the child seated appropriately in front of the computer, the experimenter asked the parent to arrange their camera view in a way that the child would not see herself or himself during testing. The parent was also asked to move the camera shot of the experimenter to the middle of the empty screen. Then, we administered a 4-point calibration designed for children in which a bouncing shape (a green star and yellow sun) appeared in four different locations on the screen. They were asked to show with their hands and arms when it appeared in each location. Calibration was repeated until the child understood how to communicate online and until the experimenter was confident in inferring the child's hand and arm movements. Once the calibration was successful, we presented each item, having 4 pictures at each corner of the screen as in Figure 4a. In every item, the experimenter presented a word, and the child was asked to choose which picture best refers to that word. Children's responses were deduced by their hand or arm movements by the experimenter. The experimenter could communicate with the child in the case of miscommunication as "*Is this the one you just showed?*" (see Figure

4b). With this design, a possible miscommunication, as can often take place in face-to-face communication, could be solved as well. Two people coded 20% of the total participants' responses for reliability analysis (Cohen's kappa = .89). Once disagreements were solved via discussion, they shared coding the rest of data.

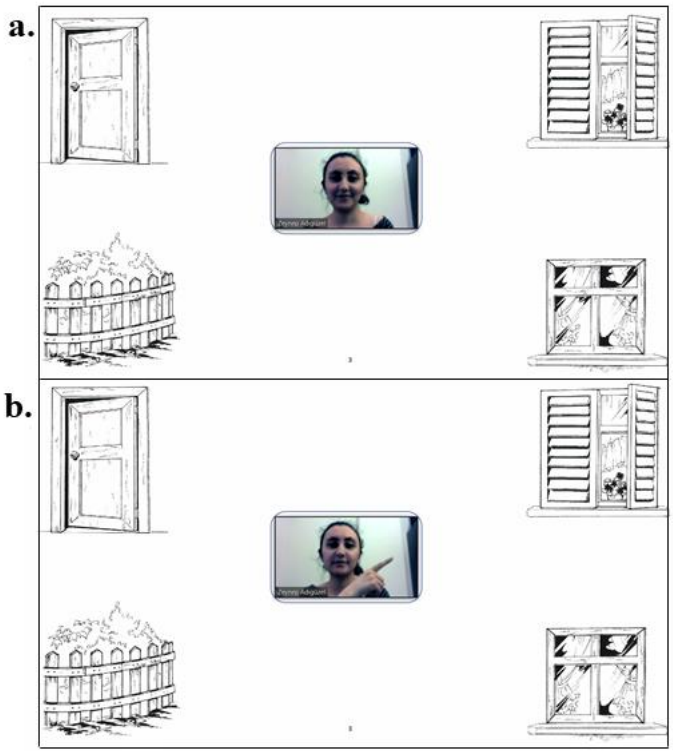

Figure 4. a) how the items were presented on the screen b) how the experimenter referred to items when needed

2. 3 Procedure

The entire procedure was implemented online via Zoom. For the decision-making task, after pilot testing, we have decided to use the Pavlovia platform to host our experiment online with no desynchronization problems between visual and auditory stimuli. Parents were instructed to open the Pavlovia link sent to them and shared their screen with the experimenter. Participants completed the tasks in two or three sessions. Each session took approximately 40-45 minutes per participant. At the beginning of the first session, parent(s) were presented with brief information regarding the study and recording conditions. Once they stated that they were ready to be recorded, the experimenter started recording and took

the verbal consent for participation. At the beginning of the session, children were warmed up with the experimenter by explaining that they were going to play a game together.

First, the decision-making task was administered. The experimenter said that this game was played by hands and positioned the child in front of their computer in the most appropriate position with the help of the parent. Children were asked to decide as if they sat side by side with the experimenter. Importantly, hand and arm movements of the child being visible was crucial so that the experimenter could understand which option was chosen. We avoided repeating the questions since it could affect children's responses (Krahenbühl & Blades, 2006). Repeating the question differently could also lead to changes in intonation and create biases over the options. Second, children were administered the working memory tasks: digit span tasks and word span tasks. Finally, the Turkish Receptive Language Test (TIFALDI-R) was administered. For half of the participants, the word span task was administered as the last task and for the other half TIFALDI-R was administered as the last task. The entire procedure was recorded via Zoom for later transcription and coding.

CHAPTER 3: RESULTS

3.1 Preliminary Analyses

First, to check whether there was a difference in two sets of questions (i.e., 2 different imaginary characters for which children were asked) in terms of the outcome variable, we conducted a dependent samples t-test. Results indicated no significant difference (p = .69). Thus, we collapsed data across the outcome variable for any further analysis.

To check whether the sex of the participants affected their responses to different sets of questions, we conducted a 2 (sex) X 2 (question set) mixed-design ANOVA. Results showed that both main effects and the interaction between variables were not significant (ps > .05). Additionally, to check whether there were sex differences in the working memory and receptive language measures, we conducted several independent samples t-tests. Results showed no significant sex difference (ps > .05). Therefore, we did not include sex as a variable in further analyses. We included age as a continuous variable (in months) in our further analyses. Given that both the forward and backward digit span tasks were not associated with children's recency score, ps > .05, we did not use the digit span tasks in our further analyses as WM measures. Figure 5 shows the relationship between age and recency score. Descriptive statistics for age, WM, language, and recency measures can be found in Table 1. Table 2 depicts the correlation matrix.

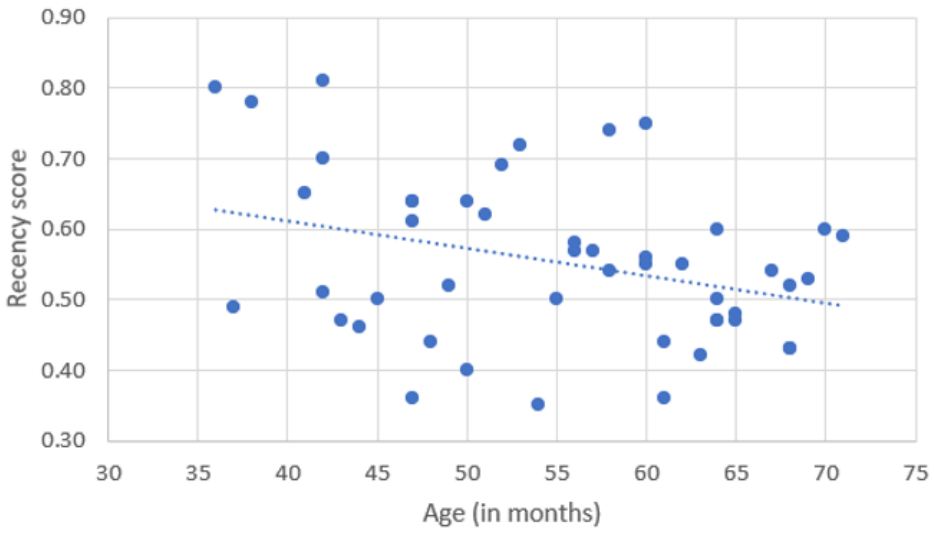

Figure 5. Scatter plot for the relationship between recency score and age (in months)

Table 1. Descriptive statistics of all measures

	N	Mean	Median	SD	Range
Age	59	53.37	53.00	10.11	36 - 71
Forward digit span	59	4.88	5.00	1.75	0 - 9
Backward digit span	59	2.14	1.00	2.28	0 - 6
Forward word span	49	5.57	6.00	1.35	2 - 9
Backward word span	49	3.10	4.00	2.28	0 - 7
Receptive language	51	67.76	72.00	17.14	20 - 92
Recency score (BB cond)	48	0.53	0.50	0.14	0.23 - 0.81
Recency score (EE cond)	48	0.55	0.53	0.15	0.20 - 0.87
Recency score (BE cond)	48	0.57	0.59	0.16	0.26 - 0.87
Recency score (average)	48	0.55	0.54	0.12	0.35 - 0.81

Note: SD = Standard deviation.
Note: BB cond: Beginning-beginning condition; EE cond: End-End condition

Table 2. Correlations between different measures

	1	2	3	4	5	6	7
1. Age	1						
2. Forward digit span score	.554**	1					
3. Backward digit span score	.647**	.514**	1				
4. Forward word span score	.581**	.735**	.518**	1			
5. Backward word span score	.658**	.610**	.745**	.634**	1		
6. Receptive language	.707**	.422**	.686**	.640**	.680**	1	
7. Recency score	-.335*	-0.274	-0.165	-.429**	-.479**	-.330*	1

*$p < .05$
**$p < .01$.

3.2 Do children exhibit recency bias after all?

Overall, 73.16% of the responses were only non-verbal responses, 1.05% were only verbal, and 25.80% were both verbal-nonverbal responses. We did not differentiate these response categories in terms of recency bias. First, we investigated whether children exhibit recency bias in their responses with one sample t-tests. Given that recency score is from 0 (no second choices were chosen) to 1 (always second choices were chosen), we tested their performance against the chance level, 0.50 to see whether they performed above chance level,

i.e., whether they show a significant tendency towards choosing the second choices). The chance level analyses showed that children's recency bias was above chance level in End-End condition (M=.55, SD= .15), $t(47)$= 2.39, p< .05, and Beginning-End condition (M=.57, SD= .16), $t(47)$= 3.01, p< .05, but not in Beginning-Beginning condition (p> .05)

To further investigate whether children showed a response bias not only towards the second option but also the first one, we first computed primacy score by subtracting recency score from 1. We then compared children's primacy and recency scores in both question sets separately and sets combined. Results of the Wilcoxon signed-rank test showed that recency score was significantly higher than primacy score in the first block (z= -2.67, p < .01), the second block (z = -2.19, p < .05), and blocks combined (z= -2.72, p < .01).

3.3 What are the underlying mechanisms of recency bias?

For our main analyses regarding underlying mechanisms of recency bias in children, we used multi-level mixed effect models (Bates et al., 2015) on R (R Core Team, 2013). Since our outcome variable was binary, we used logistic version of the model. The analyses were conducted with the lme4 package (Bates et al., 2014). We scaled our continuous predictor variables intending to prevent convergence problems in models and ease the interpretation (Bates et al., 2019). We used the *probe_interaction* function of the *interactions* package (Long, 2019) when test and plot interactions involving a continuous predictor variable. For visualization, we used the *interact_plot* function of the *ggplot2* package (Wickham, 2016).

We built two separate models, for question properties and individual differences in cognitive skills. For each model, we started with the null model (i.e., the random intercepts of participants and items as random effects). We added each predictor variable consecutively to the model and compared it with the previous model.

3. 4 Recency bias and related question properties

We first built a model to test our hypotheses about how particular question properties relate to children's preferences (Model 1). In this model, the outcome variable was preference (binary). Our fixed effects included the main effects of age and familiarity, the two-way interaction between age and familiarity, the question type, and the three-way interaction among age, familiarity, and question type. Our final sample included 50 participants and 48 trials for each participant, 2400 in total. However, there were missing observations. For this model, we selected observations that were missing on the outcome variable (preference) and excluded them on a trial-basis, which corresponds to 109 observations. Thus, our final data for this model included 2291 observations in total.

We first added age to the model. There was a main effect of age, $\chi 2$ (1) = 3.73, p = .05. One point increase in age was associated with a decrease in recency bias by .18 ± .09, z = -1.96, p < .05 We then included familiarity to the model. The inclusion of familiarity did not improve the model significantly, $\chi 2$ (1) = 1.44, p = .23. Next, we added the two-way interaction between age and familiarity, which did not significantly improve the model, $\chi 2$ (1) = 1.89, p = .17. When we added the question type to the model, no significant improvement in the model was observed, $\chi 2$ (2) = 1.73, p = .42. As the last step, we included the three-way interaction between age, familiarity, and question type. The three-way interaction did not improve the model, either $\chi 2$ (6) = 7.47, p = .28. However, slope analyses revealed that age was associated with a decrease in recency bias by .47 ± .14 for the beginning-beginning questions when the options are unfamiliar, z= -3.36, p < .05 (see Figure 6). A summary of the final model can be found in Table 3[3].

In sum, these analyses showed that children's recency bias was weaker as the age increased. Although the overall three-way interaction was not significant, slope analyses

[3] Alternatively, we also tried to build this model in a different way, by adding the main effects first, then the two-way interaction between age and familiarity, and last, the three-way interaction between age, familiarity, and question type. We obtain the same results as the previous model.

showed that the association between age and familiarity yielded a decrease in recency bias only in the Beginning-Beginning condition (i.e., when both options were at the beginning "*Elmayı mı üzümü mü saklamayı tercih eder Aymi?"*).

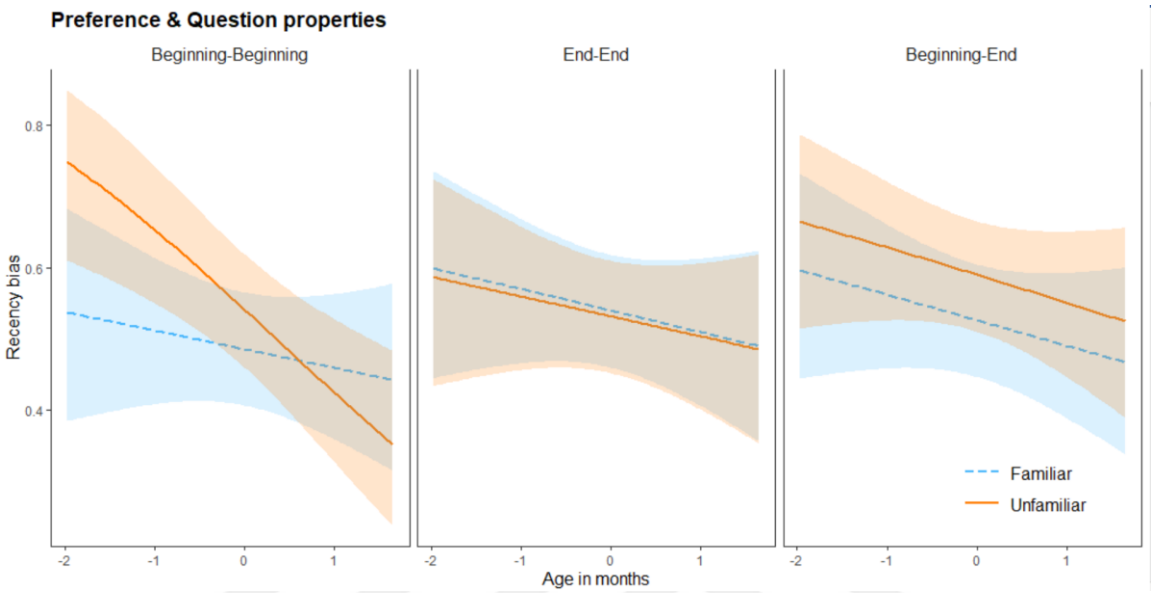

Figure 6. The relationship between recency bias (i.e., choosing the second option) and scaled scores of age in months as a factor of familiarity of options (Familiar vs Unfamiliar)

Table 3. Summary for recency bias and question properties

	Model 1	
	Coefficient	**SE**
FIXED EFFECTS		
Intercept (=FAM / Beg-beg)	-.05	.16
Age	-.10	.13
Familiarity (UNFAM)	.22	.19
Age X Familiarity	-.36*	.15
End-End Q. Type	.21	.19
Beg-end Q. Type	.16	.19
Age X Familiarity X End-End Q. Type	.37†	.21
Age X Familiarity X Beg-End Q. Type	.34	.21
	Variance	**SD**

RANDOM EFFECTS		
Subject Intercept	.34	.58
Item Intercept	.06	.25

Note: Significance codes = ***p < .001, **p < .01, *p < .05, †p < .1.

3.5 Recency bias and individual differences in cognitive skills

We built another model to examine how individual differences are in children's cognitive abilities relate to their recency biases and whether their effect might change depending on the word order. For this model, we selected all missing observations and excluded them list-wise. Thus, our final sample for this model included 1875 observations. Our fixed effects were age, the Backward Word Span Task (BWS) scores, question type, the 2-way interaction between the Backward Word Span scores and question type, and Tifaldi-R scores.

For the WM skills, we had two measures: The Digit Span Task and Word Span Task, each having both the forward and backward versions. After excluding the Digit Span Tasks due to not being associated with the outcome measure, we had the Forward and Backward Word Span tasks. Yet, these two measures were correlated with each other (r= .634, p < .01). The inclusion of both word span scores would induce collinearity. To avoid violating the assumptions of regression, we only included the scaled scores of the Backward Word Span task as the WM measure. However, our decision to pick the Backward Word Span task was not only due to statistical reasons. In the decision-making task, participants went through information encoding, maintenance, manipulation (when needed), and response implementation processes, which is similar to the central executive component of WM. This component is generally measured by complex WM tasks (e.g., backward span tasks) that require not only storage but also manipulation of information (Blom et al., 2014; Gathercole et al., 2004; Geier et al. 2009). Moreover, it is also argued that working memory capacity was heavily determined by individual differences in the functioning of the central executive

(Engle, 2002; Engle & Kane, 2004), which theoretically supports our decision to use the Backward Word Span Task as a WM measure in this mixed effect model.

We first added age to the model. There was a main effect of age, $\chi 2$ (1) = 3.85, p < .05. One point increase in age was associated with a decrease in recency bias by .18 ± .09, z= -2.01, p< .05. We then included BWS scores to the model, which improved the model at a marginal significant level, $\chi 2$ (1) = 3.62, p = .06. One point increase in BWS scores was associated with a decrease in the recency bias by .21 ± .11, z= -1.94, p = .05. The inclusion of question type, $\chi 2$ (2) = 1.9, p = .38, and Tifaldi-R scores, $\chi 2$ (1) = .71, p = .34, did not significantly improve the model, consecutively. In the last step, we included the two-way interaction between BWS scores and question type, which did not improve the model, $\chi 2$ (2) = 3.05, p = .22. Although the two-way interaction did not improve the model, our slope analyses revealed that the BWS scores were associated with a decrease in recency bias only in the beginning-beginning condition by .28 ± .14, z = -2.08, p = .04[4]. A summary of the final model can be found in Table 4.

In sum, children's WM skills (BWS) were associated with decreased recency bias, even beyond the effect of age. Also, although the overall three-way interaction was not significant, slope analyses showed that the way BWS scores associated with recency bias depended on the question type. Children's BWS scores were associated with decrease in recency bias only in beginning-beginning condition (i.e., when both options are at the beginning "*Elmayı mı üzümü mü saklamayı tercih eder Aymi?*") (see Figure 7).

[4] The recency score was significantly associated with both the Forward (r= .43) and Backward Word Span Task scores (r= .48) similarly, ps<.01. However, when we repeated the same analysis with the Forward Word Span Task scores as WM measure, we did not find significant main effect and any interaction effect including the Forward Word Span Task. See discussion section for further information.

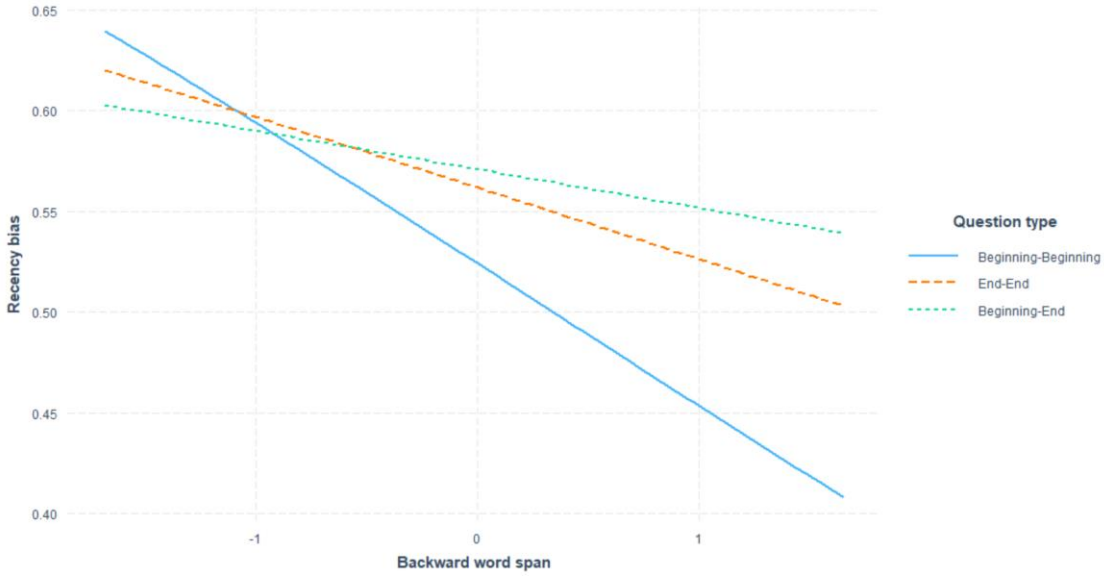

Figure 7. The relationship between recency bias (i.e., choosing the second option) and scaled scores of Backward Word Span task scores as a factor of the question type.

Table 4. Summary for recency bias and individual differences in cognitive skills

	Model 2	
	Coefficient	**SE**
FIXED EFFECTS		
Intercept (=Beg-beg)	.09	.12
Age	.01	.12
Backward Word Span Score	.28*	.13
End-End Question type	.15	.15
Beg-End Question type	.18	.14
End-End X Backward Word Span	.14	.11
Beg-End X Backward Word Span	.20 †	.12
	Variance	**SD**
RANDOM EFFECTS		
Subject Intercept	.20	.44
Item Intercept	.06	.25

Note: Significance codes = ***p < .001, **p < .01, *p < .05, †p < .1.

CHAPTER 4: DISCUSSION

The current study examined whether (a) Turkish-learning preschoolers exhibited recency bias in a flexible word order language, and (b) this bias was dependent on children's working memory and language skills. We found that children's recency bias existed and grew weaker as they age. The combination of certain question properties influenced their response tendencies. Children's WM skills, but not language skills, seemed to be a predictor of recency bias, even after controlling for age.

4.1 Children exhibit recency bias

We overall found that children exhibit recency bias when responding to two-alternative forced-choice questions. Moreover, their bias was consistent when the options were in the opposite order, i.e., they were given two different question sets (Aymi and Tibi) in which the options were in the opposite order. For example, if the question in one set was "Elmayı mı saklamayı tercih eder Aymi, üzümü mü?", it was asked as "Üzümü mü saklamayı tercih eder Tibi, elmayı mı?" in the other set. We found that children exhibit recency bias in both question sets. This suggests that, regardless of the order we presented the options, they consistently selected the last option, beyond stimulus effect.

We replicated age-related differences, showing that older children had less recency bias than younger ones. Previous studies have consistently shown that children of different ages do not behave the same way in terms of their recency tendencies in that younger children show more substantial bias than older children (Mehrani & Peterson, 2015, 2017; Sumner et al., 2019). For example, Mehrani and Peterson (2015) found that 3-year-olds had the strongest recency bias while 5-year-olds had the weakest. Behzadnia and Mehrani (2018)'s study that involved 5- and 6-years-old children showed that older preschoolers were less susceptible to recency bias than younger ones. Age-related differences might be due to still-developing cognitive skills and inexperience regarding interview-like situations. As

children age, they might become more experienced with the demands of questioning situations (Fritzley et al., 2013).

4.2 The effect of question properties on recency bias

The characteristics of the questions also seemed to be related to children's recency bias. We varied our questions in the familiarity of the objects in choices and the position of these choices within the questions. Regarding the familiarity of the options, we found that children's recency bias was not affected by the option familiarity. This was contrary to most of the previous evidence, suggesting that unfamiliar options yielded more prominent recency bias (e.g., Behzadnia & Petreson, 2017, Mehrani & Peterson, 2017). We argue that the difference between our findings and earlier research could be related to the operationalization of object familiarity. Previous studies used unfamiliar options as those exist in real life (Sumner et al., 2019, Okanda et al., 2012; Behzadnia & Mehrani, 2018) but were highly likely to be unknown by children (Fritzley, 2000). We, on the other hand, used completely unfamiliar and novel objects constructed for the present study. It is worth noting that even using unfamiliar options as of the other previous studies, Sumner et al. (2019) did not find an effect, either, possibly because of conducting the study with 2-year-olds.

Unfamiliar options may influence children's recency bias not more than, at least similarly, the familiar ones due to one other mechanism. In the unfamiliar condition, the first option was highly engaging and attention-grabbing to children, creating a surprise effect. This may have caused the first option to be better encoded, which could lead to having less difficulty in retrieving it later at the end of the sentence. Then, the relative availability or saliency of the last option would have been reduced. This way, it might be that we observed less pronounced recency bias in the unfamiliar condition than otherwise would be with unfamiliar options as of previous studies. Indeed, the same surprise effect may have increased the chances of the last option to be better encoded as well. However, since it would

not be as surprising as the first option, we may have still observed less pronounced recency bias than otherwise we would expected.

Furthermore, with this enhanced interest upon the first unfamiliar option, children's cognitive resources throughout the rest of the sentence might have become more available than they otherwise would be with the familiar options or unfamiliar options as of others. This is in line with research showing that children attend novel objects during verb learning (Behren, 1990; Golinkoff et al., 1992; Imai et al., 2008). Thus, children may have thought our novel objects were worthy of attention and treated them as truly novel. Consistent with this view, participants in the present study often asked what is it that they hear when they were presented with novel options during testing. Nevertheless, given that the main effect of object familiarity was attenuated when novel objects were used as unfamiliar objects, our results demonstrate that more work is needed to understand how the degree to which familiarity of options relate to recency bias. This can be done by experimentally manipulating familiarity on a spectrum.

We also did not find a difference between familiar and unfamiliar condition as children age. Previous evidence suggests that younger children have biases both for when options are familiar and unfamiliar whereas older children have a more pronounced bias when the options are unfamiliar (Mehrani & Peterson, 2017; Behzadnia & Mehrani, 2018). Given the attenuated effect of familiarity, the increasing benefit of cognitive maturation comes with increased age might have disappeared. In terms of the position of choices, we found no effect of the question type on recency bias. Although this is in line with Sumner et al. (2019), we expected to find the strongest recency bias in the beginning-end condition (e.g., *"**Elmayı** mı saklamayı tercih eder Aymi **üzümü** mü?").* The recency effect in this condition may have been masked by another characteristic of flexible word order languages. Although different word alternations are possible in languages like Turkish, what is near to

the main verb becomes more salient and thus, possibly, more memorable. Therefore, the tendency to choose the last option might have been challenged and faded away. The privilege of the verb-adjacent option might be underlined by children not being able to master language skills yet since they depend heavily on the syntactic relations more than the overall semantic meaning when comprehending speech. In sum, we expected the recency effect to exist in this condition due to the greater cognitive effort in retrieving the first option. However, given the saliency of the verb-adjacent option, the recency effect may have been masked or attenuated.

When we tested the interaction among conditions of familiarity, question type, and children's age, even though the overall three-way interaction was not significant, the slope analyses revealed that children's recency bias decreased significantly when responding to the unfamiliar options as they were older only in the beginning-beginning condition. As children were younger, we observe the highest recency bias in the beginning-beginning condition when options are unfamiliar (see Figure 6). The last option's being adjacent to the main verb may have contributed to recency bias to be the highest. However, as children aged, the surprise effect of unfamiliar options might have been doubled when encountered both options in a row in the beginning-beginning of the question. As the interaction was not significant, we need to cautiously interpret these slope analyses and further studies should investigate question type in more detail.

4.3 The role of individual differences in WM and verbal skills of children

In line with previous research, we found that working memory skills are related to children's recency bias. Children with better WM skills were less susceptible to recency bias than children with lower WM skills. WM had an effect on recency bias, even after controlling for age. The role of WM in children's recency tendencies is also evident in previous evidence on WM's key role in holding the earlier information until to-be-integrated information is received. Then, WM integrates it with the upcoming information later in the sentence as the

speaker constructs ideas and transforms them into the spoken language (Cowan et al., 2005, Just & Carpenter, 1992).

Our results support Sumner et al. (2019)'s conclusion that WM-based constraints are underlying factors of children's recency bias. They found that children's WM did not predict recency bias, which might be due to their WM measure not being appropriate to the sample's age. However, their participants showed stronger bias when options were longer, which led them to suggest WM-based mechanisms might be at play and to urge further research. Building upon these suggestions, our study empirically showed that children's recency bias depended on their WM capacity, particularly backward word-span task, with age-appropriate WM measures.

In terms of the association between children's WM capacity and question type, although the two-way interaction was not significant, the slope analyses again suggested a significant decrease in recency bias as WM capacity increases only in the beginning-beginning condition. As WM become more available, children should become better at holding and integrating information as the sentence unfolds its meaning. Given that WM develops late, throughout childhood into adulthood (Conklin et al., 2007; Gathercole et al., 2004), children's still developing WM skills may have not sufficed once the two options were not adjacent to each other. As the distance between two options and the distance between the onset and end of the sentence increases, children use their WM resources more. This is in line with studies on adult level comprehension as well. For example, more verbal WM resources are needed when the sentence is too long (Giustolisi et al., 2018) or when there are long-distance subject-verb dependencies within a sentence, such as in garden-path sentences ("The boat floated down the river sank" (Bever 1970). When the subject is uttered before the relative clause, it has to be hold in memory until its predicate uttered (after the relative

clause) (Giustolisi et al., 2018; Tan et al., 2017). However, we need to again approach these results with caution as these are related to slope analyses.

Moreover, WM skills are crucial for language comprehension, learning and academic success (Alloway, 2009) and problem-solving (Just & Carpenter, 1992), and cognitive control (Baddeley, 1992; Graf et al., 2002). The role of WM in sentence comprehension may also interact with the age of children (Montgomery, Evans & Gillam, 2009). As children grow older, they may be able to benefit more from the flexible processing resources of the central executive to improve the limited capacity of the phonological loop and the visuospatial sketchpad systems. This can happen either by enhancing interdependent associations between the central executive and the other sub-systems of WM. Alternatively, younger children may depend only on the phonological loop and the visuospatial sketchpad, while a dominant and salient central executive existing in older children (Gathercole et al., 2004). It is also possible that, in terms of recency bias, children might not overcome WM-based limitations in a linear fashion. Their WM capacity might remain steady and rise after reaching a certain point.

In terms of the association between working memory capacity and comprehension, it is suggested that both simple and complex spans have a role (Robertson & Joanisse, 2010; Swanson, 1992), the latter tapping more strongly to executively loaded tasks including both storage and manipulation of information. In our study, we found that both forward and backward word span scores were negatively associated with children's recency bias. However, the backward version of the task was better at detangling individual differences. This is supported by evidence showing that working memory capacity was best reflected by individual differences in the performance of the central executive (Engle, 2002; Engle & Kane, 2004).

We, however, did not find an association between children's receptive language scores and recency bias. Although verbal skills were argued to have an underlying role on

recency bias (Moriguchi et al., 2008; Okanda & Itakura, 2010), Sumner et al. (2019) did not find such an effect either. In their study, independent coders rated children's verbal skills on a scale of 1 to 4 based on their observations in the waiting room before testing. In our study, we used a standardized task, the Turkish Receptive Language Test (TIFALDI-R) (Berument & Güven, 2010), to assess children's language skills. However, since ours was an online study, the task might have been too demanding or tiring for children. Moreover, our language test measures the vocabulary knowledge. Recency bias might be associated with language skills beyond vocabulary such as syntactic knowledge. If we had used a measure requiring knowledge on the syntactic components of language, we would have observed such an effect.

4.4 Language and memory systems in general

Our study provides broader insights into language and memory processing in children. Imagine a child responding to the following questions: *"Would you like to play with a bicycle or motorcycle?".* First, from a psycholinguistic perspective, the child must comprehend the syntax and semantics of the question to answer it (Lee & Eskritt, 1999). However, this may not always be the case. Our findings challenge the assumption that children's verbal utterances are strong indicator of their word comprehension. This poses a caution especially for using parents' reports about child language development. For example, in developmental studies, parent's reports on children's verbal production were often used to indicate children's lexical knowledge with the MacArthur Bates Communicative Development Inventories (Fenson et al., 2007). If children can respond to their parents with response biases without parents being aware of it, parents can, with the same assumption, unintentionally misinform researchers (Sumner et al., 2019). In sum, the fact children can respond and act as active communication partners may not reflect their conceptual understanding. It is also possible that they may not be fully mindful decision-makers.

Second, from a situational/pragmatics perspective, the child must engage in social and communicative conventions when answering a question. Children appreciate the communicative intentions of their conversation partners early on (Liebal & Tomasello, 2009). For example, they can understand forced-choice questions requiring a response by 17 months of age (Choi, 1991). When children are asked a question, several communicative factors come into play. For example, she or he needs to keep track of factors related to the context, question, and the purpose of the communication (Mehrani, 2010). Therefore, recency bias, in a way, might reflect children's understanding that a response is being requested. Then, with this communicative urge, their effort/motivation to be a cooperative communication partner can be more pronounced than getting the specific/complete meaning. In other words, children can follow the urge to communicate without fully understanding the question itself. Understanding exact mechanisms of children's response biases may offer a window into the pragmatic development and the way in interferes with linguistic semantics. Nevertheless, how recency bias is related to a variety of social interactions remains to be an open question.

Finally, the child needs to draw on their knowledge repertoire and deliver her or his response in the form of language (Mehrani, 2010). In sum, responding to a question properly involves using knowledge at multiple levels of language processing. The existence of recency tendencies may contribute to our understanding of language processing from a variety of perspectives.

Moreover, our decision-making task can be considered a memory-dependent task. The degree to which verbal resources are recruited might depend on task's demands on memory systems. This might explain why we did not find an effect of children's verbal skills scores given that our task depends more heavily on memory systems rather than language systems. In our study, verbal resources may not respond to the task demands as strongly as working memory or memory systems did. Additionally, half of the questions children were questioned

about novel objects that cannot be sourced back to pre-existing language knowledge about the world, making language resources dispensable but memory resources indispensable. Hence, we can speculate that children's verbal skills could have had a more salient role in their decision-making processes if this had been more of a language-dependent task. The research by Kartalkanat and Göksun (2020) supported this view. They found that children's recall performance in a story-telling activity depended on their verbal abilities but not working memory. Children who had better language scores recalled more information than children who had lower language scores (Kartalkanat & Göksun, 2020). In sum, these may suggest that individual differences in memory and language systems might be associated with different processes, in other words, recruiting different resources.

4.5 Practical implications

The present study contributes to the online experiment movement in psychology (Hartshorne, 2020). Online research has a great potential to move the field forward towards a more robust science when especially accompanied with practices such as pre-registration and data sharing (Leshin et al., 2020). Recent developmental studies suggest that web-based data collection is a viable alternative to lab experiments (e.g., Nussenbaum et al., 2020; Scott & Schulz, 2017; Scott, Chu & Schulz, 2017). Our study builds on them by showing that it can be implemented with using broadly available tools (e.g., Powerpoint, Zoom) without advanced programming knowledge.

Why is it important to uncover the mechanisms governing/underlying children's response biases? Questioning children is one of the most used methods to gather information from children (Ceci & Bruck, 1993) and used in a variety of contexts including but not limited to parent-child interactions (Cameron-Faulkner et al. 2003), medical settings (von Bayer et al., 2009), eyewitness testimony and forensics (Krahenbühl & Blades, 2005; Newcombe & Dour, 2001), and social work (McGough and Warren, 1994. In addition,

response biases are important for other areas such as learning (Mavrou et al. 2013), logical reasoning (de Chantal & Markovits, 2017), and problem-solving (Cuccio-Schirripa & Steiner, 2000). Consequently, disclosing response biases children might have towards these kinds of questions holds a vital importance.

Our results revealed that when working with young children, two-alternative forces-choice questions should be used with caution, if possible, be avoided especially as children are younger. Evidence suggests that open-ended or free-recall questions can elicit longer and more accurate responses from children (Brown et al., 2013; La Rooy et al., 2015) (but see also Peterson et al., 1999; Waterman et al., 2001). However, they may fail to respond appropriately when questions are unclear and incomprehensible (Fritzley, 2000; Peterson et al., 1999), and especially younger ones may do so given their underdeveloped cognitive capabilities (Aldridge & Wood, 1998; Fritzley et al., 2013). In situations where forced-choice questions are required (i.e., children are not giving information as much as needed otherwise), researchers should avoid repeating questions by rephrasing them and make sure alternative choices are fairly counterbalanced. We would also recommend that, as we benefited a great deal during data collection, adult questioners should minimize environmental/external distractions, take some time to build a warm connection with the child before questioning and explain what the child exactly should expect next.

Investigating children's response tendencies has implications not only for applied settings but also for developmental research (Mehrani, 2010). While forensic interviews generally rely on unstructured questions to gather as much information as accurately as possible, developmental research relies more heavily on structured questions so that the variation among children cannot be due to the differences in questions per se (Fritzley, 2000). The way children are questioned as an empirical method is often neglected in developmental research (Ceci & Bruck, 2006; Fritzley et al., 2013) although questioning children has been

the primary data collection method in empirical developmental studies (Behzadnia & Rad, 2020; Fritzley & Lee, 2003). Without better understanding biases leaking into developmental research, we may never reveal their true capabilities or biases to which they suffer from (Fritzley et al., 2013). Therefore, it is crucial to understand how developmental phenomena are tested and how questioning procedures affected the participants (Fritzley, 2000). Understanding how children respond to questions will potentially lead the studies to have more appropriate designs to gather information from children (Behzadnia & Rad, 2020).Our results extend this research by showing that not only question-related factors, but also individual differences WM capacity may be linked to children's decision-making processes.

4.6 Limitations and future directions

The current study provides a systematic investigation of children's recency bias when answering two-alternative forced-choice questions. One of the strengths of our work is that it draws advantage from the mixed effect analyses that allow us to simultaneously take all potential factors into consideration (Baayen, Davidson & Bates, 2008). Still, it has some limitations. One caveat of the present study is about sample characteristics. In fact, with online data collection, we could reach a sample relatively more diverse than a lab sample with respect to geography (i.e., families from outside of İstanbul, where the lab is situated, could join the project, such as Adana, Eskişehir, Mersin, Ankara, Antalya). The sample was still a convenience sample. It involved families who heard about the project via our advertisements and could afford a computer and stable internet connection. Second, although we could achieve certain amount of variability in our individual difference tasks, future studies can use, for example, targeted advertising to further increase the sample diversity. This can also be effective for particular future projects that needs to reach out certain populations that would otherwise be beyond reach (Leshin et al., 2020).

Third, online data collection can bring new challenges for researchers working with children. Since our whole procedure required participants being required to be in front of a computer, parents had to sit aside the entire time. We made sure that parents sit behind their children and not limited their movements. We also discarded trials where parents attempted to invade children's decision-making processes. However, parent presence might still influence children. Moreover, we also aimed to control the size of all visuals by requiring our participants to join the zoom call via computer rather than any other device. Unfortunately, it was impossible to have high control over it as in laboratory testing since everyone has different technological equipment. More work is required to better delineate the possible situations that can occur in online testing as well as how appropriate precautions can be improved.

Although being in the comfort of their own homes with their accustomed devices, children may have found online testing less engaging compared to laboratory settings. We tried to control the prosody and intonation in the decision-making task by playing pre-recorded audio. Even though the experimenter was present and could be seen by participants during testing, children still knew that they were not interacting completely with the experimenter. Children may have been more attentive if the experimenter always communicated live. However, this can also be a limitation since it may increase the chance of having differences in stress among trials and options within trials. We were also careful in selecting the verbs to be commonly used and neutral. Future research can further investigate whether recency bias would change depending on the concreteness, frequency and valence (positive vs negative) of the actions in which children were questioned about. In addition, the online implementation of the measure of individual differences in receptive language (Tifaldi-R) was probably more difficult -and potentially more problematic than laboratory counterpart- because it took the longest time among our tasks. Therefore, children could

become bored more in less time. Future studies can address these issues by employing shorter tasks and different paradigms including a more active participation from children.

CHAPTER FIVE: CONCLUSION

In conclusion, this study examined children's recency tendencies in a free word order language with respect to individual differences in a systematic way. It is important to note that the present research was first to investigate children's recency tendencies to forced-choice questions in a flexible word order language. Children exhibited recency bias even when the options were switched between each other, which suggests a true difficulty in their decision-making processes. Older children experienced weaker recency bias than younger children. Whether the options were familiar and positioned in different word orders did not have an effect in isolation. Older children had weaker recency bias than younger ones when responding to unfamiliar options that both were positioned at the beginning of the question. Children with higher WM skills (but not verbal skills) were less susceptible to recency bias. Individual differences in cognitive resources available to children seemed to interact with word order alternations, such that, children's recency bias decreased as their WM skills increased when both options were encountered at the beginning of the question. In addition, the descriptive information regarding each word order shows that children had the strongest bias in beginning-end condition. This is not necessarily contradicting with the findings of multi-level mixed-effect models. Rather, it suggests that children exhibit recency bias that interacts with individual differences in cognitive skills.

Together, our findings shed light on children's preferences, suggesting that it has foundations in heuristics and biases, especially early in life. Response biases are possibly valuable in making quick decisions and judgements at the earlier stages of language and memory development. However, as children age and their cognitive skills develop, they overcome this dependency and become freer of their biases.

REFERENCES

Acarlar, F., Aksu-Koç, A., Küntay, A. C., Maviş, İ., Sofu, H., Topbaş, S., & Turan, F. (2008, August). Adapting MB-CDI to Turkish: The first phase. In *Essays of Turkish Linguistics: Proceedings of the 14th International Conference on Turkish Linguistics*, 6-8.

Akoglu, G., & Acarlar, F. (2014). Investigation of Turkish nonword repetition list for 3-9 years children. *Egitim ve Bilim*, *39*(173), 13-24.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable?. *Child Development*, *77*(6), 1698-1716. https://doi.org/10.1111/j.1467-8624.2006.00968.x

Anderson, V. A., Anderson, P., Northam, E., Jacobs, R., & Catroppa, C. (2001). Development of executive functions through late childhood and adolescence in an Australian sample. *Developmental Neuropsychology*, *20*(1), 385–406. https://doi.org/10.1207/S15326942DN2001_5.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language*, *59*(4), 390-412.

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation*, *8*, 47–89. Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. Retrieved from http://arxiv.org/abs/1406.5823

Behrend, D. A. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child development*, *61*(3), 681-696.

Behzadnia, A., & Mehrani, M. B. (2018). Persian-Turkish bilingual children's responses to forced-choice questions. *Current Psychology, 1*(7). doi:10.1007/s12144-018-9814-x

Berthier, N., DeBlois, S., Poirier, C., Novak, M., & Clifton, R. (2000). Where's the ball? Two- and three-year-olds reason about unseen events. *Developmental Psychology, 36*(3), 394–401. https://doi.org/10.1037//0012-1649.36.3.394

Blake, J., Austin, W., Cannon, M., Lisus, A., & Vaughan, A. (1994). The relationship between memory span and measures of imitative and spontaneous language complexity in preschool children. *International Journal of Behavioral Development, 17*(1), 91–107. doi:10.1177/016502549401700106

Booth, J. R., Mac Whinney, B., & Harasaki, Y. (2000). Developmental differences in visual and auditory processing of complex sentences. *Child Development, 71*(4), 981-1003. DOI: https://doi.org/10.1111/1467-8624.00203

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Boyle, W., Lindell, A. K., & Kidd, E. (2013). Investigating the role of verbal working memory in young children's sentence comprehension. *Language Learning, 63*(2), 211-242.

Brady, M. S., Poole, D. A., Warren, A. R., & Jones, H. R. (1999). Young children's responses to yes-no questions: Patterns and problems. *Applied Developmental Science, 3*(1), 47-57. https://doi.org/10.1207/s1532480xads0301_6

Ceci, S. J., & Bruck, M. (2006). Children's suggestibility: Characteristics and mechanisms. *Advances in Child Development and Behavior, 34*, 247-281. https://doi.org/10.1016/S0065-2407(06)80009-1

Conklin, H. M., Luciana, M., Hooper, C. J., & Yarger, R. S. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental Neuropsychology*, *31*(1), 103–128. https://doi.org/10.1207/s15326942dn3101_6

Cowan, N. (2005). Working memory capacity limits in a theoretical context. In C. Izawa, N. Ohta (Eds.), *Human learning and memory: Advances in theory and application. The 4th Tsukuba international conference on memory* (pp. 155-175). Mahwah, NJ: Erlbaum. https://doi.org/10.4324/9781410612717-12

Cowan, N., Elliott, E. M., Scott Saults, J., Morey, C. C., Mattox, S., Hismjatullina, A., & Conway, A. R. A. (2005). On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, *51*(1), 42–100. https://doi.org/10.1016/j.cogpsych.2004.12.001

Daviesl, G. M., Westcott, H. L., & Horan, N. (2000). The impact of questioning style on the content of investigative interviews with suspected child sexual abuse victims. *Psychology, Crime and Law*, *6*(2), 81-97. DOI: 10.1080/10683160008410834

Emmorey, K., Giezen, M. R., Petrich, J. A. F., Spurgeon, E., & O'Grady Farnady, L. (2017). The relation between working memory and language comprehension in signers and speakers. *Acta Psychologica*, *177*, 69–77, https://doi.org/10.1016/j.actpsy.2017.04.014

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19-23.

Engle, R. W., & Kane, M. J. (2004). Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory,* Vol. 44, pp. 145–199). Elsevier Science.

Fenson L, Bates E, Dale PS, Marchman VA, Reznick JS, Thal DJ. MacArthur-Bates communicative development inventories. Paul H. Brookes Publishing Company; 2007.

Fritzley, V., Lindsay, R. C., & Lee, K. (2013). Young children's response tendencies toward yes–no questions concerning actions. *Child Development*, *84*(2), 711-725. DOI: 10.1111/cdev.12006

Fritzley, V., & Lee, K. (2003). Do young children always say yes to yes-no questions? A metadevelopmental study of the affirmation bias. *Child Development*, *74*(5), 1297–1313. https://doi.org/10.1111/1467-8624.00608

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177. DOI: 10.1037/0012-1649.40.2.177

Geier, C. F., Garver, K., Terwilliger, R., & Luna, B. (2009). Development of working memory maintenance. *Journal of neurophysiology*, 101(1), 84-99.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental psychology*, *28*(1), 99.

Hartshorne, J. K. (2020). How massive online experiments (MOEs) can illuminate critical and sensitive periods in development. *Current Opinion in Behavioral Sciences*, *36*, 135-143.

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child development*, *79*(4), 979-1000.

Jun Zhang, L., & Bin Anual, S. (2008). The role of vocabulary in reading comprehension: The case of secondary school students learning English in Singapore. *RELC Journal*, *39*(1), 51–76. https://doi.org/10.1177/0033688208091140

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149. http://dx.doi.org/10.1037/0033-295X.99.1.122

Just, M. A., Carpenter, P. A., & Keller, T. A. (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, *103,* 773–780. doi: 10.1037/0033- 295X.103.4.773

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of experimental psychology: General*, *133*(2), 189.

Kartalkanat, H., & Göksun, T. (2020). The effects of observing different gestures during storytelling on the recall of path and event information in 5-year-olds and adults. *Journal of experimental child psychology*, *189*, 104725.

Kersten, A. W. (1998). A division of labor between nouns and verbs in the representation of motion. *Journal of Experimental Psychology: General*, *127*(1), 34.

Kersten, A. W., & Smith, L. B. (2002). Attention to novel objects during verb learning. *Child development*, *73*(1), 93-109.

Kim, J. (2020). Learning and teaching online during Covid-19: Experiences of student teachers in an early childhood education practicum. *International Journal of Early Childhood*, *52*(2), 145-158.

Kwon, H., Reiss, A. L., & Menon, V. (2002). Neural basis of protracted developmental changes in visuo-spatial working memory. *Proceedings of the National Academy of Sciences*, *99*(20), 13336–13341. https://doi.org/10.1073/pnas.162486399

Leshin, R. A., Leslie, S. J., & Rhodes, M. (2020). Does It Matter How We Speak About Social Kinds? A Large, Preregistered, Online Experimental Study of How Language Shapes the Development of Essentialist Beliefs. *Child development*. *92*(4), 531-547

Liebal, K., & Tomasello, M. (2009). Infants appreciate the social intention behind a pointing gesture: Commentary on "Children's understanding of communicative intentions in the middle of the second year of life" by T. Aureli, P. Perucchini and M. Genco. *Cognitive Development*, *24*(1), 13–15. doi:10.1016/j.cogdev.2008.09.004

Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, *76*(3), 697–712. https://doi.org/10.1111/j.1467-8624.2005.00872.x

Malatesha Joshi, R. (2005). Vocabulary: A critical component of comprehension. *Reading & Writing Quarterly*, *21*(3), 209–219. https://doi.org/10.1080/10573560590949278

Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading*, *31*(3), 273–284. https://doi.org/10.1111/j.1467-9817.2008.00371.x

McDuffie, A. S., Yoder, P. J., & Stone, W. L. (2006). Labels increase attention to novel objects in children with autism and comprehension-matched children with typical development. *Autism*, *10*(3), 288-301.

Mehrani, M. B., & Peterson, C. (2015). Recency Tendency: Responses to forced-choice questions: Recency tendency. *Applied Cognitive Psychology*, *29*(3), 418–424. https://doi.org/10.1002/acp.3119

Mehrani, M. B., & Peterson, C. (2017). Children's recency tendency: A cross-linguistic study of Persian, Kurdish and English. *First Language*, *37*(4), 350–367. https://doi.org/10.1177/0142723717694055

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology*, *41*(1), 49-100.

Moghadam, S. H., Zainal, Z., & Ghaderpour, M. (2012). A review on the important role of vocabulary knowledge in reading comprehension performance. *Procedia - Social and Behavioral Sciences*, *66*(1), 555–563. https://doi.org/10.1016/j.sbspro.2012.11.300

Montgomery, J. W., Evans, J. L., & Gillam, R. B. (2009). Relation of auditory attention and complex sentence comprehension in children with specific language impairment: A preliminary study. *Applied Psycholinguistics*, *30*(1), 123-151.

Moriguchi, Y., Okanda, M., & Itakura, S. (2008). Young children's yes bias: How does it relate to verbal ability, inhibitory control, and theory of mind? *First Language*, *28*(4), 431–442. https://doi.org/10.1177/0142723708092413

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Nussenbaum, K., Scheuplein, M., Phaneuf, C. V., Evans, M. D., & Hartley, C. A. (2020). Moving developmental research online: Comparing in-lab and web-based studies of model-based reinforcement learning. *Collabra: Psychology*, *6*(1).

Okanda, M., & Itakura, S. (2010). When do children exhibit a "Yes" bias? *Child Development*, *81*(2), 568–580. https://doi.org/10.1111/j.1467-8624.2009.01416.x

Okanda, M., Kanda, T., Ishiguro, H., & Itakura, S. (2013). Three-and 4-year-old children's response tendencies to various interviewers. J*ournal of Experimental Child Psychology*, *116*(1), 68-77. http://dx.doi.org/10.1016/j.jecp.2013.03.012

Okanda, M., Somogyi, E., & Itakura, S. (2012). Differences in response bias among younger and older preschoolers: Investigating Japanese and Hungarian preschoolers. *Journal of Cross-Cultural Psychology*, *43*(8), 1325-1338. http://dx.doi.org/10.1177/0022022112440145

Peterson, C., & Grant, M. (2001). Forced-choice: Are forensic interviewers asking the right questions?. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences du Comportement*, *33*(2), 118-127. https://psycnet.apa.org/doi/10.1037/h0087134

Pickering, S., & Gathercole, S. E. (2001). *Working Memory Test Battery for Children (WMTB-C) manual*. London, UK: Psychological Corporation.

Raman, I., Raman, E., & Mertan, B. (2014). A standardized set of 260 pictures for Turkish: Norms of name and image agreement, age of acquisition, visual complexity, and conceptual familiarity. *Behavior research methods*, *46*(2), 588-595. https://doi.org/10.3758/s13428-013-0376-4

Savaşır, I., & Şahin, N. (1995). Wechsler çocuklar için zeka ölçeğii (WISC-R) el kitabı. Ankara, Turkey: Türk Psikologlar Derneği Yayınları.

Scott, K., & Schulz, L. (2017). Lookit (part 1): A new online platform for developmental research. *Open Mind*, *1*(1), 4-14.

Scott, K., Chu, J., & Schulz, L. (2017). Lookit (Part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15-29.

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The syllable repetition task (SRT). *Journal of Speech, Language, and Hearing Research : JSLHR*, *52*(5), 1189–1212. https://doi.org/10.1044/1092-4388(2009/08-0047)

Steffensen, M. S. (1978). Satisfying inquisitive adults: Some simple methods of answering yes/no questions. *Journal of Child Language*, *5*(2), 221–236. https://doi.org/10.1017/S0305000900007431

Sumner, E., DeAngelis, E., Hyatt, M., Goodman, N., & Kidd, C. (2015). Toddlers always get the last word: Recency biases in early verbal behavior. In *Proceedings of the Thirty-Seventh Annual Conference of the Cognitive Science Society*

Sumner, E., DeAngelis, E., Hyatt, M., Goodman, N., & Kidd, C. (2019). Cake or broccoli? Recency biases children's verbal responses. *PLOS ONE*, *14*(6), e0217207. https://doi.org/10.1371/journal.pone.0217207

Swanson, H. L. (1992). Generality and modifiability of working memory among skilled and less skilled readers. *Journal of Educational Psychology*, *84*(4), 473-488. https://psycnet.apa.org/doi/10.1037/0022-0663.84.4.473

Tomasello, M. (2001). Perceiving intentions and learning words in the second year of life. In M. Tomasello & E. Bates (Eds.), *Language development: The essential readings*. 111–128. Blackwell Publishing.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*(6), 635-654.

Valenti, R., Staiano, J., Sebe, N., & Gevers, T. (2009). Webcam-based visual gaze estimation. In *International Conference on Image Analysis and Processing*, 662-671. Springer, Berlin, Heidelberg.

APPENDICES

Appendix A: List of objects and questions used across conditions

		Choice 1	Choice 2	Question
both choices at the beginning	Familiar	Kapı	Masa	**Kapıyı** mı **masayı** mı yıkamayı tercih eder Aymi?
		Ayakkabı	Ceket	**Ayakkabıya** mı **cekete** mi dokunmayı tercih eder Aymi?
		Tren	Uçak	**Treni** mi **uçağı** mı kurcalamayı tercih eder Aymi?
		Bıçak	Kaşık	**Bıçağı** mı **kaşığı** mı kullanmayı tercih eder Aymi?
	Unfamiliar	Burer	Nılha	**Bureri** mi **nılhayı** mı kurcalamayı tercih eder Aymi?
		Değa	Unve	**Değayı** mı **unveyi** mi kullanmayı tercih eder Aymi?
		Büma	Lerun	**Bümayı** mı **lerunu** mu yıkamayı tercih eder Aymi?
		Pızadeni	Terka	**Pızadeniye** mi **terkaya** mı dokunmayı tercih eder Aymi?
both choices at the end	Familiar	Lamba	Televizyon	Hangisini sallamayı tercih eder Aymi, **lambayı** mı **televizyonu** mu?
		Kulak	Parmak	Hangisini tutmayı tercih eder Aymi, **kulağı** mı **parmağı** mı?
		Araba	Bisiklet	Hangisini yüzdürmeyi tercih eder Aymi, **arabayı** mı **bisikleti** mi
		Kuş	Tavuk	Hangisine bakmayı tercih eder Aymi, **kuşa** mu **tavuğa** mı?
	Unfamiliar	Dedinubi	Santu	Hangisini sallamayı tercih eder Aymi, **dedinubiyi** mi **santuyu** mu?
		Dör	Kudar	Hangisine bakmayı tercih eder Aymi, **döre** mi **kudara** mı?
		Dakül	Zasır	Hangisini tutmayı tercih eder Aymi, **dakülü** mü **zasırı** mı?
		İlmasan	Kobika	Hangisini yüzdürmeyi tercih eder Aymi, **ilmasanı** mı **kobikayı** mı?
one choice at the beginning one at the end	Familiar	Elma	Üzüm	**Elmayı** mı saklamayı tercih eder Aymi **üzümü** mü?
		Kamyon	Otobüs	**Kamyonu** mu almayı tercih eder Aymi, **otobüsü** mü?
		Böcek	Karınca	**Böceği** mi yakalamayı tercih eder Aymi **karıncayı** mı?
		Limon	Soğan	**Limonu** mu yıkamayı tercih eder Aymi **soğanı** mı?
	Unfamiliar	Yalem	Yedar	**Yalemi** mi saklamayı tercih eder Aymi **yedarı** mı?
		Bikirum	Maka	**Bikirumu** mu yakalamayı tercih eder Aymi **makayı** mı?
		Buriç	Cukil	**Burici** mi yıkamayı tercih eder Aymi **cukili** mi?
		Kenu	Nesanbaş	**Kenuyu** mu almayı tercih eder Aymi, **nesanbaşı** mı?

Note: List for the first block. In the second block, the order of choices were flipped for each question.

Appendix B: The Full List of Image Stimuli

The entire image stimuli for familiar and unfamiliar choices can be found in this link: https://docs.google.com/document/d/1BFdthDOG_e7v2Dw5Yr2gyXf0MwVJUOAeyEs4VLRMMPE/edit

Appendix C: The Forward and Backward Word Span Tasks

a. Forward Word Span Task

Deneme 1: Fıstık – Havlu **Deneme 2: Sakız- Çiçek**

1 hece	Fil-tuz	Çöp-Top
2 hece	Yangın -Şeker	Baykuş – Fırın
3 hece	Süpürge- Şemsiye- Telefon	Tuvalet- Çaydanlık – Pencere
4 hece	Battaniye- Matematik- Kalorifer- Televizyon	Bilgisayar –Kaplumbağa- Muhallebi- Ayakkabı

b. Backward Word Span Task

Deneme1: Market- Köprü **Deneme 2: Bahçe- Perde**

1 hece	Tuz-Saç	Bot-süt
2 hece	Reçel-şapka	Kukla – Tarak
3 hece	Patates – Sinema – Kumanda	Tarhana – Hastane zürafa
4 hece	Salatalık- Helikopter- Karasinek- Uyurgezer	Kertenkele- Denizaltı- Doğumgünü- Kütüphane