DEVELOPMENT OF EVENT REPRESENTATION IN LANGUAGE AND COGNITION: EVIDENCE FROM TURKISH SPEAKING CHILDREN

BİLGE TINAZ

ÖZYEĞİN UNIVERSITY

MAY, 2024

DEVELOPMENT OF EVENT REPRESENTATION IN LANGUAGE AND COGNITION: EVIDENCE FROM TURKISH SPEAKING CHILDREN

by
Bilge Tınaz

Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Arts

in
Psychology

Advisor: Assoc. Prof. Ercenur Ünal

Graduate School of Social Sciences
Özyeğin University
İstanbul

May, 2024

DEVELOPMENT OF EVENT REPRESENTATION IN LANGUAGE AND COGNITION: EVIDENCE FROM TURKISH SPEAKING CHILDREN

Approved by:

Assoc. Prof. Ercenur Ünal, Advisor

Department of Psychology

Özyeğin University

Asst. Prof. Alper Açık

Department of Psychology

Özyeğin University

Assoc. Prof. Aslı Aktan Erciyes

Department of Psychology

Kadir Has University

Approval Date: 29.05.2024

DECLARATION OF ORIGINALITY

I hereby declare that I am the sole author of this thesis and that this is the true copy of my thesis, including the final revisions, approved by my thesis committee. All data and information have been obtained, produced, and presented in accordance with the rules of research ethics and principles of academic honesty. As required by these rules, to the best of my knowledge I have acknowledged ideas, thoughts, and any copyrighted material in accordance with the standard referencing rules. I certify that any part of this thesis has not been submitted for a degree or diploma in another educational institution.

Bilge Tınaz

ABSTRACT

During language development, children need to make sense of a continuous stream of input into meaningful segments, known as events, and communicate about them. Here, we examine how preschoolers (4- and 5-year-olds) and adults segment events in language and cognition in the domain of motion events, focusing on Turkish, a verb-framed language that segments motion paths in separate verb phrases. In the non-linguistic Dwell Time task, participants viewed self-paced slideshows created with frames taken from videos of motion events that either involved a path change or did not involve a path change. Detection of an event unit is indicated by event boundaries, where the time spent looking at a slide is significantly longer compared to other slides. In the linguistic task, the same participants described the same events to a confederate addressee. Our findings revealed that participants described motion events that involve a path change with two or more path verbs more frequently than events that do not. We observed a developmental pattern in the linguistic descriptions: adults tended to use multiple verb phrases more frequently than younger children and older children tended to use them more frequently than their younger counterparts. This tendency was particularly evident in using two or more path verbs. However, despite the differences in linguistic descriptions, children and adults segmented the events similarly in the non-linguistic task. These findings suggest a lack of homology between cognitive and linguistic event segmentation throughout development.

Keywords: Motion events, event segmentation, dwell time, eye-tracking

ÖZET

Dil gelişimi sürecinde, çocuklar çevrelerinde meydana gelen olayları anlamlı birimlere bölmeyi ve bunlar hakkında iletişim kurmayı öğrenmelidir. Bu çalışmada, Türkçe'ye odaklanarak dil ve biliş alanında çocukların ve yetişkinlerin hareket olaylarını nasıl birimlere ayırdığı incelenmektedir. Türkçe, hareketin yönünü ayrı fiillerle ifade eden fiil odaklı bir dildir. Sözel olmayan görevde katılımcılar yön değişimi içeren veya içermeyen hareket olaylarının videolarından alınan ekran görüntüleriyle oluşturulmuş slayt gösterilerini izlediler. Katılımcılar bir tuşa basarak bir sonraki resme geçtiler, yani resimlerin ekranda kalma süresini kendileri belirlediler. Bir olay sınırı, bir slayta bakılan sürenin diğer slaytlara göre anlamlı ölçüde daha uzun olmasıyla belirlendi. Sözel görevde aynı katılımcılar aynı olayları bir sahte katılımcıya anlattılar. Bulgularımız, katılımcıların yön değişikliği içeren hareket olaylarını, içermeyen olaylara göre daha sık iki veya daha fazla yol fiili ile anlattıklarını gösterdi. Sözel görevde bir gelişimsel farklılık görüldü: yetişkinler ve 5 yaş çocuklar olayları anlatırken 4 yaş çocuklardan daha sık olarak birden fazla fiil kullandılar. Bu farklılık, iki veya daha fazla yön fiili kullanımında özellikle belirgindi; yetişkinler, 5 yaşındakilere kıyasla birden fazla yön fiili kullanmaya daha meyilliydiler, 4 yaşındakiler ise bunları nadiren kullandılar, bu da henüz dile özgü yapıları tam olarak edinmediklerini göstermektedir. Ancak, sözel görevdeki farklılıklara rağmen, hem çocuklar hem de yetişkinler sözel olmayan görevde olayları benzer şekilde birimlere ayırdılar. Bu bulgular, olayları birimlere ayırmada gelişim boyunca bilişsel ve dilsel süreçlerin birbirinden bağımsız olduğunu göstermektedir.

Anahtar Kelimeler: Hareket olayları, olayları birimlere ayırma, bakma süresi, göz izleme

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Ercenur Ünal, for her unwavering guidance, invaluable insights, continuous support, and dedication to ensuring the rigor and quality throughout the entirety of this work. Her mentorship has not only shaped this thesis but has also equipped me with valuable research skills that I will carry forward in my academic career. I extend my heartfelt appreciation to the members of my thesis committee, Dr. Alper Açık and Dr. Aslı Aktan-Erciyes, for their time and constructive feedback.

Thanks to TÜBİTAK, the Scientific and Technological Research Council of Turkey for the 2210-A Master's Fellowship that has provided me with the financial support necessary to pursue my academic aspirations. This study has been supported by a TÜBİTAK 1001 grant, awarded by the Scientific and Technological Research Council of Turkey (TÜBİTAK) (grant number 121K259) to Dr. Ercenur Ünal.

I am grateful to all members of the Language, Cognition, and Development Lab, particularly Ceren, Bahar, Teoman, Aleyna, Sarp, Eren, Dursunay, Lena, and Elif, for their invaluable assistance in data collection and coding, which significantly contributed to the richness of the findings presented in this thesis. A special appreciation goes to Ceren, who contributed significantly to this work by assisting in stimulus preparation and providing invaluable support throughout the entire process. Having Ceren and Bahar as friends made the ups and downs of writing this thesis much more enjoyable. I would like to express my sincere gratitude to all the participants and children who participated in this study.

I would like to express my deepest love and thanks to my friends Betül, Gökçe, Kübra, and Sercan, whose encouragement, understanding, and companionship have provided solace during the challenging moments and made the journey of completing this thesis more enjoyable.

I owe a debt of gratitude to Ege, for his enduring love, encouragement, and patience throughout this academic journey. His unwavering support has been a constant source of strength and inspiration.

Last but certainly not least, I want to express my profound gratitude to my family for their unconditional love, belief in me, and endless sacrifices. Their support and encouragement have been the cornerstone of my academic pursuits. Finally, I dedicate my thesis to my nephew, Arda, who was born during the process of writing the thesis.

TABLE OF CONTENTS

DECLARATION OF ORIGINALITY iii
ABSTRACT iv
ÖZET v
ACKNOWLEDGMENTS vi
LIST OF TABLES x
LIST OF FIGURES xi
1. INTRODUCTION 1
    1.1 The Relation Between Language and Cognition 1
    1.2 Event Segmentation in Cognition 6
    1.3 Event Segmentation in Language 9
        1.3.1 Acquisition of Language-specific Motion Expressions 12
    1.4 Linguistic Influences on Event Segmentation 15
    1.5 The Current Study 17
2. METHOD 21
    2.1 Participants 21
    2.2 Materials 22
        2.2.1 Norming Study 23
    2.3 Procedure 25
        2.3.1 Linguistic Task 25
        2.3.2 Non-linguistic Dwell Time Task 26
    2.4 Coding 27
    2.5 Preprocessing of the Eye Data in the Dwell Time Task 30
3. RESULTS 31
    3.1 Linguistic Task 31
        3.1.1 Analysis of More Than One Verb Phrase 31
        3.1.2 Analysis of Two or More Path Verbs 33
    3.2 Non-linguistic Dwell Time Task 35
4. DISCUSSION 39
    4.1 Event Segmentation in Language and Cognition 39
    4.2 Developmental Patterns in Event Segmentation 42

4.3 Theoretical Implications 44
4.4 Methodological and Practical Implications 47
4.5 Limitations and Future Directions 48
4.6 Conclusion 50
REFERENCES 52
APPENDICES 62
APPENDIX A. FIXED EFFECT ESTIMATES FOR GLMER MODEL OF MULTICLAUSE USE 63
APPENDIX B. FIXED EFFECT ESTIMATES FOR GLMER MODEL OF MULTIPLE PATH CLAUSE USE 64
APPENDIX C. PROPORTION OF MULTICLAUSE USE 65
APPENDIX D. PROPORTION OF MULTIPLE PATH CLAUSE USE 66
APPENDIX E. FIXED EFFECT ESTIMATES FOR GLMER MODEL OF BUTTON PRESSES 67

LIST OF TABLES

Table 2.1 List of Target Events Used in the Experiment 22

Table 2.2 Path and Manner Verbs Used by Different Age Groups................................ 28

Table 3.1 Parameter Estimates for the Fixed Effects from the Glmer Model for Using More Than One Verb Phrase 33

Table 3.2 Parameter Estimates for the Fixed Effects from the Glmer Model for Using Two or More Path Verbs 35

Table 3.3 Parameter Estimates for the Fixed Effects from the Lmer Model for the Mean Fixation Durations 37

Table 3.4 Parameter Estimates for the Fixed Effects from the Glmer Model for Using Multiclause.. 63

Table 3.5 Parameter Estimates for the Fixed Effects from the Glmer Model for the Using Two or More Path Clause.. 64

Table 3.6 Parameter Estimates for the Fixed Effects from the Lmer Model for the Mean Latency of the Button Presses... 67

LIST OF FIGURES

Figure 2.1 Distribution of Weekly English Hours of Child Participants 21

Figure 2.2 Examples of Boundary (Outlined in Red) and Non-Boundary Slides: (A) Path-Change (B) No Path-Change 24

Figure 3.1 Proportion of More Than One Verb Phrase Across Event Types 31

Figure 3.2 Proportion of Descriptions With at Least Two Path Verbs Across Event Types 34

Figure 3.3 Mean Fixation Durations Across Event Types and Slides 36

Figure 3.4 Proportion of Multiclause Across Event Types 65

Figure 3.5 Proportion of Descriptions With at Least Two Path Clauses Across Event Types 66

1. INTRODUCTION

During language development, children need to learn to organize a continuous stream of activity into meaningful units, known as events, and communicate about them. Events can be perceived and organized in a hierarchical manner, with coarse units subsuming finer units. This hierarchical organization involves the segmentation of events in different levels of granularity (Zacks et al., 2001). For example, we might construe a trip to a coffee shop as one large event unit or as a series of finer units such as walking into the coffee shop, approaching the counter, ordering the coffee, finding a seat, drinking the coffee, and leaving the coffee shop. However, events are inherently complex, both cognitively and linguistically, as each comprises several components, and world languages vary in expressing information about them. Consequently, understanding how language-cognition interactions in event segmentation change during development becomes a fundamental question. Despite this, no study has comprehensively examined the development of event segmentation using both linguistic and non-linguistic methods with participants spanning various age groups. Hence, the current study aims to address this gap by gathering cognitive, linguistic, and developmental perspectives. We ask two research questions: (1) Do linguistic differences in event segmentation influence non-linguistic event segmentation in cognition? (2) Do linguistic influences on event segmentation change developmentally throughout language acquisition?

1.1 The Relation Between Language and Cognition

Understanding the complex relationship between language and cognition has long been a primary objective of cognitive science, with various theories positing different perspectives on the nature of this relationship. According to the universalist view, language acquisition mostly maps onto a pre-linguistic foundation since humans have a conceptual framework for representing people, objects, and their spatial relationships (Gleitman & Papafragou, 2016; Landau et al., 2010). Conversely, the relativistic view contends that since language and conceptions are primarily co-constructed, the representation of concepts varies cross-linguistically (Levinson, 2003; Majid et al., 2004).

According to the results of many cross-linguistic studies with adults, speakers of different languages show similar performances in non-verbal tasks such as visual attention or memory, even though they express the same events differently in their language (Cardini, 2010; Feinmann, 2020; Gennari et al., 2002; Papafragou et al., 2008; Trueswell & Papafragou, 2010).

A study by Choi and Bowerman (1991) explored how English and Korean learning infants showed sensitivity to language-specific lexicalization patterns when describing motion events. English-learning infants tended to focus on the manner (how an action is done) of motion from an early age, incorporating a wide range of verbs into their descriptions of events. Korean-learning infants, conversely, were more likely to emphasize the path (the trajectory of an action) of motion in their early expressions, reflecting the lexicalization patterns of their native language. The researchers argued that linguistic representations significantly shape speakers' perception of motion events, suggesting that children's spatial learning is influenced by language-specific structures, rather than directly mapping spatial words onto pre-existing nonlinguistic concepts. However, this study focused only on the lexicalization patterns to investigate how children acquire spatial concepts. Furthermore, a study by Pruden and colleagues (2004) investigated how infants categorize path and manner in motion events. They found that by 10 months old, infants could distinguish the path, and by 15 months, they could distinguish the manner. In another study by Pulverman and colleagues (2008), it was found that 14-to-17-month-old infants from both English- and Spanish-speaking communities can distinguish changes in manner and path in motion events. These findings suggest that infants possess the cognitive ability to categorize aspects of events before developing language skills, indicating foundational cognitive abilities to understand events before verbalizing them. These studies indicate that elements guiding event cognition adhere to universal principles arising from general cognitive tendencies. However, these studies made inferences about conceptual development based solely on pre-linguistic perception and did not independently examine how children expressed events later in development or how they are conceptualized after language acquisition. Furthermore, a study conducted with English- and Japanese-learning infants discovered that although infants start with language-general event categories, they become more sensitive to the components of motion events expressed in their native language from the age of 19 months onwards, suggesting a shift in event conceptualization as children are exposed to a specific language (Göksun et al., 2011). In this study, while the researchers investigated how events are conceptualized at various stages of language acquisition, linguistic expressions were not investigated. Specifically, exploring how infants' discrimination of event components relates to the development of specific language structures was not taken into account. Furthermore, a recent longitudinal study examined how Turkish-learning children between the ages of 1 and 3 express the spatial features of events in language, and how it relates to their conceptual skills (Aktan-Erciyes & Göksun, 2019). The study found that infants' ability to detect motion event components at one year old predicted their verb comprehension, and how they lexicalize event components specific to their native language at three years old. The findings from this research indicate a connection between early event processing and verb comprehension, as well as how children express these elements in accordance with the linguistic demands of their native language. However, in this study, universal tendencies were present in the linguistic descriptions of children at the age of 3.

There are also a limited number of studies that examined language and cognition interactions in the motion event domain cross-linguistically and developmentally (Bunger et al., 2021; Papafragou et al., 2002; Papafragou & Selimis, 2010). For example, Papafragou and colleagues (2002) investigated how English and Greek speakers describe motion events, as well as how linguistic differences influence the performances on non-linguistic tasks such as categorization and memory performance. Despite language-specific patterns in motion event descriptions, English and Greek speakers performed similarly in non-linguistic tasks. The study found that within language linguistic descriptions did not consistently correlate with memory and categorization performance. Furthermore, regardless of language, adults more frequently used manner verbs in describing events compared to children. Specifically, English adults showed a greater tendency to use manner verbs than English children, whereas Greek adults and children showed similar patterns in using manner verbs. Interestingly, despite these age group variations in linguistic descriptions, both English and Greek speakers, whether children or adults, demonstrated similar categorization preferences in non-linguistic tasks. This suggests a consistency in how speakers categorize motion events across different age groups within each language group. Overall, the findings of this study suggest an independence between conceptual and linguistic representation of motion events. However, in this study, older children, such as 8-year-olds, participated in the categorization task, which suggests that they are likely to have acquired certain language-specific structures at that age.

Taken together, these studies provide valuable insights into the interaction between language and cognition. However, a critical gap exists in our understanding of the developmental continuity of this interaction. While research has explored both infants, toddlers, and adults, preschoolers remain understudied (but see Papafragou et al., 2002.) This lack of data hinders our ability to fully comprehend how language acquisition relates to event perception across different age groups. Furthermore, there is a limited number of studies that have developmentally examined the interaction between language and cognition in event perception by comparing situations where language is used versus where it is not.

In summary, studies on event cognition have revealed different perspectives on whether language influences event cognition. This thesis project delves into these varying perspectives in the domain of event segmentation with a particular focus on the developmental relationship between the segmentation of motion events in language and cognition. As detailed in section 1.3, motion events are a good test bed for testing this because there are cross-linguistic and developmental differences regarding how certain event components are expressed (e.g. Allen et al., 2007; Bohnemeyer et al., 2007; Talmy; 2000). Furthermore, as explained in section 1.2, events are spontaneously segmented in cognition when there are changes in the spatiotemporal characteristics (Zacks et al., 2007; Zacks & Tversky, 2001). Moreover, although adults mostly comply with the segmentation patterns in language, children acquiring language show both universal and language-specific tendencies in the process of language acquisition (Allen et al., 2007; Bunger et al., 2012; Papafragou et al., 2002). This duality makes this domain an ideal testing ground for studying how children acquire typological distinctions in their language and how this process relates to the development of corresponding cognitive event concepts. Thus, this study aims to examine event segmentation through both linguistic and non-linguistic measures across different age groups, including preschoolers and adults. To our knowledge, this is the first study to investigate event segmentation in language and cognition by comparing children and adults.

1.2 Event Segmentation in Cognition

According to Event Segmentation Theory (Zacks et al., 2007; see also Radvansky & Zacks, 2017) our minds naturally segment continuous experiences into meaningful events, which function as the foundational units for memory and cognition. Working memory representations of events, known as *event models*, maintain information about the core features of an event including its participants and spatiotemporal framework. When core features of the event are stable, the event model can make accurate predictions about upcoming happenings. However, when the core features of the event change, the event model cannot accurately predict what will happen next and has to be updated. This update is perceived as an event boundary. In their study, Zacks et al. (2011) investigated how prediction errors prompt changes in perception and memory and affect how people perceive events. The study involved participants watching movies of naturalistic events and making predictions about what would happen next. Consequently, unpredictable moments corresponded to subjective event boundaries in the stream of experience. These increases in prediction error serve as signals to update working memory and reorient attention to salient features in the environment, resulting in the experience of event boundaries. An event boundary is triggered when there are changes to the people or objects in the event (Zacks et al., 2009), intentionality (Baldwin et al., 2001; Saylor et al., 2007), goal-related structure (Zacks, 2004; Kosie & Baldwin, 2021), causality (Cohen & Oakes, 1993), and spatial characteristics such as direction (Zacks, 2004).

The majority of empirical evidence on event segmentation is based on explicit judgments of event boundaries. In a classical paradigm developed by Newtson (1973), participants are instructed to watch a movie showing an actor performing several activities and to indicate each new unit with a button press. Studies using this paradigm have shown that people can consistently identify the breakpoints in the event in a hierarchical manner (Sargent et al., 2013; Sasmita & Swallow, 2023; Zacks et al., 2001). Furthermore, the movie frames that correspond to the breakpoints are remembered better than the frames that correspond to within-unit moments (Newtson & Engquist, 1976). However, one caveat of this method is that it requires substantial verbal instruction and explanation regarding events and what ‘fine’ and ‘coarse’ event units might correspond to. Furthermore, event segmentation is considered to occur prior to linguistic processing (Zacks et al., 2007). Therefore, it would be beneficial to use measurements that do not contain linguistic input.

This caveat is addressed by the Dwell Time paradigm developed by Hard and colleagues (2011). The "boundary advantage" refers to the fact that people tend to pay more attention to moments that coincide with event boundaries than to moments that occur within event units (Hard et al., 2011; Kosie & Baldwin, 2019b). In accordance with this, participants are presented with a series of still images sampled from videos of continuous actions at regular intervals (Hard et al., 2011). They are asked to advance through the slide show at their own pace by pressing a button while the time between the button presses (i.e., how long they dwell on each image) is recorded. Studies using this method have shown that slides that correspond to event boundaries are viewed for a longer time compared to the slides that fall between event boundaries (Hard et al., 2011; Kosie & Baldwin, 2021; Meyer et al., 2011; Sage & Baldwin, 2014; Zheng et al., 2020). This increase in looking time is attributed to the additional demand for attention at event boundaries that is required for consolidating event representations in memory. Importantly, the looking time data from the Dwell Time task correlates with explicit boundary judgments (Hard et al., 2011).

Due to its non-verbal and implicit nature, this method has been deemed suitable for use with diverse populations, including children as young as 2.5 years old (Kosie & Baldwin, 2021; Meyer et al., 2011; Zheng et al., 2020). Meyer and colleagues (2011) investigated 3- and 4-year-olds' hierarchical event segmentation by adapting the Dwell Time paradigm. In their study, children, similar to adults in prior research (Hard et al., 2011), exhibited increased attention at event boundaries compared to within-unit moments, with variations depending on the granularity of the event. Specifically, children showed a linear trend in dwell times, with longer durations for coarse-level breakpoints and shorter durations for fine-level breakpoints. Zheng and colleagues (2020) further explored the development of event segmentation and memory in children aged 5- to 7-years-old compared to adults by using the Dwell Time paradigm. Children segmented continuous activities depicted in the self-paced slideshows into distinct events. Their segmentation patterns were consistent among different observers and in adults who performed the same task. Furthermore, similar to adults, the ability to segment events also correlated with children's memory recall of the events. Similarly, Kosie and Baldwin (2021) examined the processing of unfolding motion concerning the goal-related structure of events in 2.5- to 4.5-year-old children. Using the Dwell Time paradigm, the researchers investigated how children prioritize goal-related segmental structure over various motion patterns as events unfold. The study revealed that preschoolers exhibited selective attention at event boundaries. Like adults, they prioritized goal-related segmental structure over various motion patterns when processing unfolding events.

Overall, Event Segmentation Theory suggests that people spontaneously segment a continuous stream of activity into meaningful units, serving as the building blocks for memory and cognition. Event models encode core features of events and make predictions about upcoming happenings. Changes in these features prompt updates in perception and memory, leading to the perception of event boundaries. Research using paradigms like explicit judgments (Newtson, 1973) and the Dwell Time paradigm (Hard et al., 2011) confirms that people consistently detect event boundaries and allocate more attention to these moments. The Dwell Time paradigm is a more implicit measurement and is suited for diverse populations including young children. It has been shown that how long people dwell on an image correlates with event boundaries. Through this method, it has been shown that the duration of time people spend dwelling on an image correlates with event boundaries.

1.3 Event Segmentation in Language

Beyond perception, people frequently communicate about the events they perceive. Each event comprises several components and there are language-specific constraints regarding how much information about event components can be packaged in a single linguistic unit (i.e., a verb phrase; Levelt, 1989).

A well-attested case is motion events (Talmy, 2000; Bohnemeyer et al., 2007; see also Ünal et al., 2023). Motion events (e.g., *a ball rolled into a house*), involve a moving entity known as the figure (a ball), with respect to a landmark known as the ground (a house), along a trajectory or path (into) in a certain way or manner (rolling).

According to Talmy's (2000) typological classification, languages differ based on whether they encode the path of motion in verbs or satellites of the verb. In satellite-framed languages (e.g., English, German, Dutch) manner of motion is typically expressed in the main verb, while path is expressed through satellites (e.g., particles or adpositions). Since manner verbs can be freely combined with different path segments, satellite-framed languages can tightly package manner and path into a single clause (or verb phrase), even if there are multiple path segments (see example 1).

(1)	The	ball	rolled	down	into	the house
			verb	preposition	preposition	
			manner	path	path	

By contrast, in verb-framed languages (e.g., French, Turkish, Spanish) path of motion is typically expressed in the main verb, while the manner of motion is optionally expressed through adverbs, subordinate verbs, or adpositions. As a result, each path segment or change in the direction of motion is expressed in a new verb phrase (see example 2).

(2)	Top	yuvarlan-arak	in-di	ve	ev-e	gir-di
	ball	roll-CONN	descend-PST	and	house-DAT	enter-PST
		sub. verb	verb			verb
		manner	path			path

'The ball rollingly descended and entered the house.'

Talmy's (2000) typological classification offers a helpful frame for understanding how world languages represent events using various linguistic categories. Previous research has demonstrated that speakers of languages categorized differently by this typology indeed express motion events according to these lexicalization patterns (Bohnemeyer et al., 2007; Slobin, 1996; see also studies in Ünal et al., 2023). Similar findings have been observed across various language comparisons, such as Turkish-English (Özçalışkan, 2015; Özçalışkan & Slobin, 2003; Özyürek et al., 2008), English-Turkish-Japanese (Allen et al., 2007; Kita & Özyürek, 2003), English-Turkish-Spanish (Özçalışkan & Slobin, 1999), English-Japanese-Spanish (Maguire et al., 2010), Greek-English (Papafragou, Hulbert & Trueswell, 2008; Papafragou, Massey & Gleitman, 2002; 2006; Papafragou & Selimis, 2010), Spanish-English (Feinmann, 2020; Gennari et al., 2002; Naigles et al., 1998), French-English (Hickmann et al., 2009), and Italian-English (Cardini, 2010). This line of research indicates that the described typologies represent the predominant ways in which languages express event descriptions, distinguishing between satellite-framed and verb-framed languages.

In addition to Talmy's typological classification, Bohnemeyer and colleagues (2007) introduced the macro-event properties (MEP) as a method for measuring linguistic event segmentation by investigating 18 typologically diverse languages. MEP refers to how languages package event information, ensuring temporal operators, such as adverbs, adverbials, conjunctions, and tense markers, encompass all subevents, and present events with unique initial and/or terminal boundaries, durations, and positions. Different from Talmy's typology, which categorizes languages as verb-framed or satellite-framed based on motion event encoding, MEP focuses on how languages tightly express event-related information. According to MEP, languages are classified into three types based on their encoding strategies: Type-I languages integrate departure, passing, and arrival subevents into a single expression, achieved through satellite-framing or multiverb constructions. Type-II languages encode departure and arrival together but require separate expressions for passing subevents, often using a double-marking strategy. Type-III languages isolate each location-change subevent in separate expressions, exclusively expressing paths in verb roots. For instance, in example 1, a single macro-event expression is used while in example 2, different path segments are isolated in different verb roots. In summary, MEP offers insights into how languages segment and encode motion events, complementing Talmy's typology by focusing on the tight packaging of event information in linguistic expressions. Overall, both classifications offer a useful basis for investigating the segmentation of motion events in language. For the case of multipath events, while MEP focuses more on the packaging of event-related information, Talmy's linguistic typology focuses on whether the path component is encoded in the main verb and its consequences for clausal-level packaging.

1.3.1 Acquisition of Language-specific Motion Expressions

Building upon our understanding of how languages differ in expressing motion events, it's crucial to examine the acquisition of language-specific structures in children. Previous work has shown that children typically start describing events in a language-general manner and they progress to mapping different linguistic elements onto different components in accordance with language-specific patterns around the age of three (Allen et al., 2007; Hickmann et al., 2009; Maguire et al., 2010; Özçalışkan & Slobin, 1999; Özyürek et al., 2008; Papafragou, Massey & Gleitman, 2002; 2006; Papafragou & Selimis, 2010). For instance, Allen and colleagues (2007) examined whether 3-year-old children speaking English, Japanese, and Turkish used language-specific structures when describing motion events. Their findings revealed that English-speaking children tended to package path and manner within a single clause more frequently than their Turkish-speaking counterparts. Turkish-speaking children typically expressed path within the main verb and manner within a subordinate clause. Although children speaking Turkish and Japanese showed patterns similar to adults by mapping path and manner information across two separate clauses, Turkish-speaking children also demonstrated a higher frequency in the use of single-clause expressions compared to Turkish-speaking adults. This indicates a potential universal tendency to represent motion events in language. Another study conducted by Özyürek and colleagues (2008) found that English-speaking children used single clauses while Turkish-speaking children used two clauses to express manner and path starting from the age of 3. Comparable results were obtained from a study involving English- and French-speaking children aged between 3 and 7 (Hickmann et al., 2009). In this study, English-speaking children exhibited a more frequent inclusion of both manner and path information in their expressions compared to French-speaking children. Similar findings came from studies conducted with English- and Greek-speaking children (Bunger et al., 2021; Papafragou, Massey & Gleitman, 2002; 2006; Papafragou & Selimis, 2010). According to the findings of these studies, children between the ages of 3 and 10 who speak Greek and English, described the manner and path of the motion using structures specific to their native languages. For instance, 8-year-old Greek-speaking children were less likely to use manner verbs compared to English-speaking children (Papafragou, Massey & Gleitman, 2002). Moreover, Greek-speaking children tended to omit manner and provide path information exclusively in their descriptions of motion events.

However, universal tendencies have also been observed in some studies. For example, although English-speaking children used manner verbs more frequently than path verbs, 3-years-old Turkish- and Spanish-speaking children used path and manner verbs at similar rates when describing motion events (Özçalışkan & Slobin, 1999). Similarly, in a study by Aktan-Erciyes and Göksun (2019), Turkish-speaking 3-year-olds used manner-only motion event descriptions more frequently than path-only and path-manner descriptions. Researchers argue that these findings imply that language-specific patterns may not be fully established by the age of three, making it challenging for children of this age to integrate both path and manner into a single description. This is consistent with the findings of Allen and colleagues (2008) which suggested that universal tendencies for encoding motion may continue to exist at the age of 3. Moreover, Bunger and colleagues (2012), investigated linguistic omissions of path and manner components in children's event descriptions. They examined whether these omissions stem from developmental differences in attention allocation during event conceptualization or limitations of the language production system. They found that 4-year-olds were more likely to omit path or manner than adults in their descriptions. Overall, findings revealed similarities in event apprehension between 4-year-olds and adults, suggesting that linguistic omissions are not solely due to differences in attention. Instead, they are likely linked to the developing language production system.

Studies on how children acquire language-specific structures have revealed a complex developmental trajectory characterized by the interaction between universal tendencies and language-specific patterns. Typically, children transition from language-general descriptions of events to mapping linguistic elements onto different components in accordance with the structures of their native languages. This transition is usually observed around the age of three. Studies conducted across various languages have demonstrated this progression. Although language-specific patterns often emerge, universal tendencies still persist, suggesting ongoing developmental processes in language acquisition. The differences in language use between adults and children highlight the dynamic nature of language development.

1.4 Linguistic Influences on Event Segmentation

Could cross-linguistic differences in the segmentation of events influence event segmentation in language? Until recently, the competing views described above could not be tested in the domain of event segmentation as linguistic and non-linguistic segmentation of events has been examined independently. However, a recent study by Gerwien and von Stutterheim (2018) investigated this by comparing native speakers of French (a verb-framed language) and German (a satellite-framed language). The study examined the formation of event units in motion events that involve orientation/direction change through both linguistic and non-linguistic tasks. In the linguistic task, participants described motion events shown in brief video clips. French speakers were significantly more likely to produce more than one verb phrase to describe the motion events that had a change in orientation/direction. In contrast, German speakers were less likely to do so. In the non-linguistic task, another group of French and German speakers performed the Newtson task (1973) to segment the same motion events. That is, participants were asked to press a button when they perceived a change in the situation in the video. French speakers were more likely to indicate that there was an event boundary compared to German speakers. Specifically, French speakers' detection of the additional event boundary corresponded to the points when the figure changed direction/orientation. The cross-linguistic differences between French and German observed in both linguistic and non-linguistic tasks were taken as evidence for the presence of strong language-driven influences on cognitive event unit formation.

Although the findings reviewed above reveal a parallel between event units in language and cognition, a few aspects of the study challenge this interpretation. First, in the linguistic task, the descriptions were only coded in terms of whether or not the participant used more than one verb phrase referring to the motion events. However, the content of the phrases (path, manner, or both) was not considered. This is important because even though in verb-framed languages manner is typically encoded outside of the main verb, manner verbs can also be used (even though there are fewer manner verbs available). Thus, it might be possible to produce a description consisting of more than one verb phrase without using a new phrase to refer to a different path segment but simply by expressing manner and path—two motion components occurring simultaneously—in separate verb phrases (e.g., rolled and entered).

Second, it is possible that in the non-linguistic task, participants may have been implicitly verbalizing even though they were not explicitly required to respond verbally. This is especially likely for the study by Gerwien and von Stutterheim (2018) given that participants completed boundary judgments on the Newtson task (1973) by interpreting verbal instructions (i.e., to press a button "when [they] perceive a change/whenever something new happens"). Due to its explicit and verbal nature, this task might have encouraged linguistic encoding during non-linguistic event segmentation, even if the response was non-verbal (i.e., button press). In fact, previous work has shown that cross-linguistic differences disappear when people are prevented from implicitly using language while performing a non-linguistic task (e.g., Trueswell & Papafragou, 2010; Winaver et al., 2007), leaving open the possibility that the cross-linguistic differences reported by Gerwien and von Stutterheim (2018) reflects thinking with language rather than non-linguistic cognition.

Despite recent evidence that sheds light on how language might influence event segmentation, there are still several unanswered questions about the robustness of these linguistic influences in adults. The study by Gerwien and von Stutterheim (2018) provided valuable insights into how linguistic differences affect event segmentation, but it has some limitations. For example, the study focused only on whether participants used more than one verb phrase in describing motion events, without considering the differences in the content of these phrases. Moreover, there are concerns about whether participants implicitly verbalized during non-linguistic tasks, which raises questions about the extent to which language may have influenced the segmentation.

Moving forward, investigating how language influences event segmentation in children presents an interesting research direction to address unanswered questions. It is necessary to build upon our current knowledge of language acquisition in childhood and look into the relationship between language-specific structures and the development of event segmentation patterns. Previous studies have indicated that the gradual development toward adult-like use of linguistic structures involves a complex interaction between language and cognition that requires further explanation. Moreover, it is important to observe whether cognitive processes change with the acquisition of linguistic structures during development. By addressing these questions, we can deepen our understanding of how language and cognition interact in event segmentation.

1.5 The Current Study

In the present study, we investigate the developmental relation between linguistic and non-linguistic event segmentation. Several limitations of prior work discussed above remain regarding our understanding of how the linguistic descriptions of events relate to event segmentation at the conceptual level. Despite extensive literature on how children acquiring various languages express events (e.g., Allen et al., 2007; Hickmann et al., 2009; Maguire et al., 2010; Özçalışkan & Slobin, 1999; Özyürek et al., 2008; Papafragou, Massey & Gleitman, 2002; 2006), these studies have predominantly overlooked non-linguistic measures of event segmentation. In our study, we aim to address these open issues by examining event segmentation in language and cognition across participants from different age groups. We focus on native speakers of Turkish, a verb-framed language that allows expressing motion events in multiple clauses with separate verb phrases for each path segment. We use a within-language comparison by examining event segmentation with linguistic and non-linguistic measures within the same language population comprising both children and adults. With this approach, we endeavor to understand the nature of language-cognition interactions in event segmentation.

The current study has two aims. Our first aim is to investigate the extent to which linguistic differences in event segmentation influence non-linguistic event segmentation in cognition. Regarding this aim, we address the limitations of previous work in the following ways. First, we conduct a more detailed examination of event descriptions in language. As described in the previous section, people can distribute path and manner into different verb phrases, but since these components often occur simultaneously, expressing them with different phrases in language may not necessarily indicate that they correspond to different cognitive event units. Therefore, we delve into the content of the descriptions to ensure that descriptions with multiple verb phrases indeed convey a change in direction/orientation. Second, we use the Dwell Time paradigm (Hard et al., 2011), which offers a more implicit measure of event segmentation. To further validate this measure, we also record participants' eye movements as an index of attention allocation as they complete the Dwell Time task.

Our second aim is to examine potential developmental changes in linguistic influences on event segmentation throughout language acquisition. To do so, we included participants from three age groups in our study: 4-year-olds, 5-year-olds, and adults. The reason we selected these age groups is that, as mentioned in section 1.3.1, children aged 4 can still omit path or manner components when describing motion events (Bunger et al., 2012).

In the linguistic task, participants described videos of motion events that involved a change in the direction of motion. As a control, we also included motion events that do not involve a change in direction. Given that Turkish allows the expression of each path segment with a separate verb phrase, of interest was whether participants would be more likely to produce descriptions that consist of multiple verb phrases for events that involve a path change as opposed to those that do not have a path change. Since language-specific structures may not be fully acquired by the age of 4, we expect the production of descriptions with multiple verb phrases will increase with age. Specifically, we expect 5-year-olds to use these structures more frequently than 4-year-olds, and adults to use them more frequently than children. Also, of interest was whether events described using multiple verb phrases were indeed due to the use of two or more path verbs, especially for events that involve a change in the direction of motion.

In the non-linguistic task, participants completed an eye-tracked Dwell Time task. We examined whether there would be an increase in Dwell Time for the images that corresponded to the event boundaries (i.e., change in direction of motion).

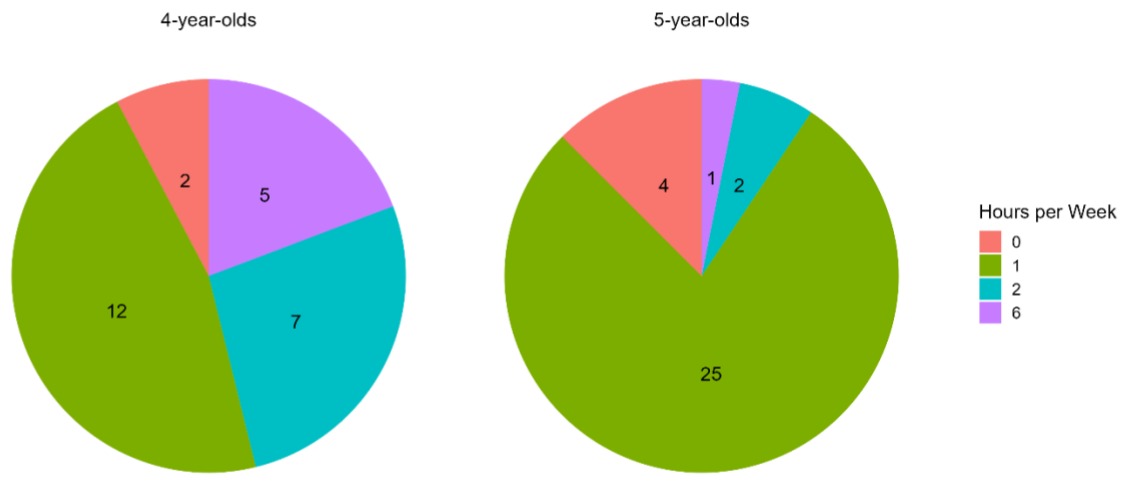
Assuming our predictions for the linguistic task are supported, the two theoretical positions reviewed above make different predictions for the non-linguistic task. If language shapes how people perceive events, participants should segment events that involve a path change vs. those that do not differently. Thus, Dwell Time should increase for the slides that depict a path change. Furthermore, this effect should be stronger in adults compared to children who may not yet acquired language-specific structures. Alternatively, if event categories are largely shared regardless of language, then the number of event units in the segmentation task may not necessarily parallel the number of event units in language. In this case, there are two plausible outcomes: participants may segment events involving a path change and those that do not into an equal number of units across all age groups, or events might be segmented into finer units in cognition before they are segmented into finer units in language. Consequently, we would expect Dwell Times to increase at event boundaries, but children might not necessarily use multiple verb phrases when describing them. Furthermore, although Dwell Times increase in path change slides in all age groups, adults should describe events more frequently with multiple verb phrases and path verbs than children, and 5-year-olds should describe events more frequently with multiple verb phrases and path verbs than 4-year-olds.

2. METHOD

2.1 Participants

Data were collected from 85 native Turkish speakers from three age groups: adults (n = 31, 20 females, M_{age} = 21.96 years, SD_{age} = 2.41, range = 18.87 – 27.84), 4-year-olds (n = 26, 18 females, M_{age} = 4.49, SD_{age} = 0.34, range = 3.85 – 4.99) and 5-year-olds (n = 32, 21 females, M_{age} = 5.58, SD_{age} = 0.32, range = 5.17 - 6.29). Children were recruited from local private and public kindergartens in İstanbul, Turkey. The medium of education in these kindergartens was Turkish. However, children were exposed to English through language classes. Figure 2.1 shows the variation in weekly English lesson hours that children receive. Adult participants were undergraduate students at Özyeğin University and received course credits for their participation. Data from 8 additional children and two additional adults were excluded because their native language was not Turkish (n = 1), they did not follow the instructions (n = 7) or they did not fit the age criteria (n = *2*).

Figure 2.1 *Distribution of Weekly English Hours of Child Participants*



Note. Numbers on the graph depicts the number of children receiving the corresponding hours of English lessons.

2.2 Materials

Target stimuli consisted of motion events depicting geometric objects changing location in a specific manner and along a trajectory with respect to a landmark object. There were two types of motion events: events that involved a path change (e.g., a ball rolled down a hill, and moved along the lateral axis to enter into a house; Figure 2.2.A) and events that did not involve a path change (e.g., a ball spun along the lateral axis to enter into a tower; Figure 2.2.B). There were also filler events that did not involve any path change and any landmark (e.g., a triangle jumping to the left) (see Table 2.1 for a complete list of events).

Table 2.1 *List of Target Events Used in the Experiment*

	Event Type	Target events
1	path-change	spin up past
2	path-change	roll down into
3	path-change	jump up past
4	no path-change	spin into
5	no path-change	roll down
6	no path-change	jump up

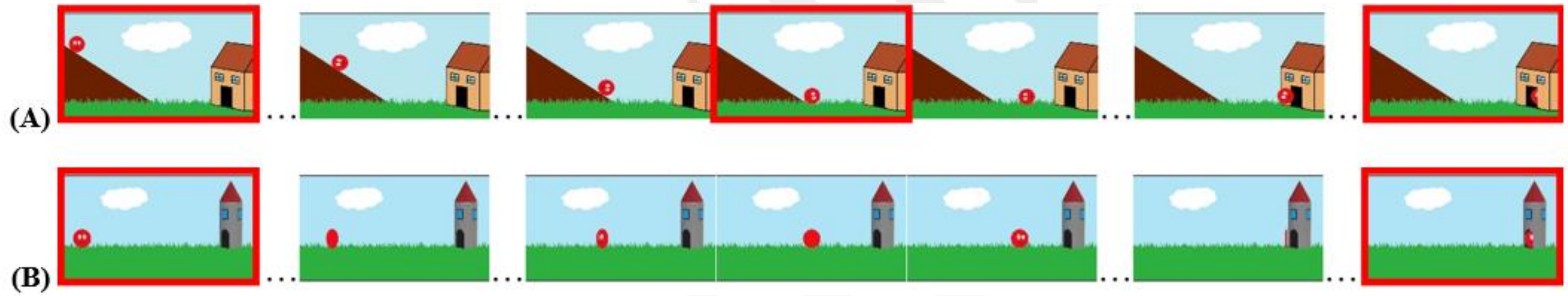
Stimuli was created in Adobe Premiere Pro CC 2015. Each video was 10 seconds long. Videos featured a sky-blue background with a cloud and green grass on the ground. Each video incorporated a landmark object, combined with a moving figure to establish distinct motion paths. For *into* paths, landmarks were positioned near the motion's endpoint. For *past* paths, landmarks were placed towards the motion's end in a way that allowed the moving object to pass them. Paths involving *upward* or *downward* motion were represented using a hill or stairs.

For the non-linguistic task, we created a slideshow by sampling screenshots from the videos at 1-second intervals. Thus, each 10-second video was converted into a slideshow consisting of 11 images (see Figure 2.2). For path-change events, the slide corresponding to the change in direction was set to be in the middle, that is, on the 6^{th} slide, for each slideshow.

2.2.1 Norming Study

Prior to the experiment, event boundaries that corresponded to path changes were independently verified in a norming study. Instructions and procedures were adapted from Kosie and Baldwin (2021) who used a similar Dwell Time task. Participants were provided with definitions of events and event boundaries, along with examples. They were instructed that an event could be viewed as a single event or as smaller sub-events, each considered a separate event. Participants (n = 30) were shown the complete sequence of slides in order and asked whether each slide depicted a boundary or not. Slides considered to be a boundary were coded as 1, and slides considered not to be a boundary were coded as 0. We computed the proportions of participants' decisions marking the path-change slide (Slide 6) and the slide leading up to the path-change (Slide 5) as boundaries. Mean boundary judgments for the path-change slide (Slide 6) and the slide leading up to the path-change (Slide 5) were significantly higher for events that involved a path change (M = 0.46) as opposed to the events that did not involve a path change (M = 0.30, χ^2 (1) = 9.238, p = .002).

Figure 2.2 *Examples of Boundary (Outlined in Red) and Non-Boundary Slides: (A) Path-Change (B) No Path-Change*



2.3 Procedure

Child participants were tested individually in a quiet room at the kindergartens. Adult participants were tested individually in the lab. They all were seated approximately 60cm away from a DELL Precision M4800 laptop with the SMI RED 250 eye-tracker (SensoMotoric Instruments) mounted underneath the screen. The stimuli were presented via NBS Presentation software (Version 23.1, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

All participants received both the linguistic and non-linguistic Dwell Time task in a fixed order: they performed the non-linguistic task first and the linguistic task second. This was done to prevent transfer from a task that involves using language to one that does not.

2.3.1 Linguistic Task

Participants watched nine video clips of motion events displayed on the computer. After watching each video clip, they described what happened in the video to a confederate addressee who was seated across them and could not see the computer display. Individual videos were presented in a single randomized order, with half of the participants receiving the original list and the remaining half receiving the items in the reversed order. Prior to the main experiment, participants completed two practice trials. In the first, a circle figure rotated in place, while in the second, a triangle figure jumped in place. Participants were instructed to watch the videos carefully and describe the events (what happened in the video) to the addressee when a gray screen appeared at the video's end. After the initial practice trial, participants were prompted to focus on the main event rather than the background elements in subsequent videos if necessary. Upon successful completion of the second trial, the main experiment started. For adult participants, after completing the practice trials, the experimenter handed the mouse to the confederate addressee to proceed with the remaining videos, leaving the participants alone in the room. In contrast, for child participants, to prevent random clicking, the experimenter sat beside the child throughout the experiment, advancing to the next video once the description was provided.

During the implementation phase, no additional differences were introduced for adult participants, except that while the descriptions of adult participants were video-recorded, those of child participants were voice-recorded due to practical reasons.

2.3.2 Non-linguistic Dwell Time Task

Participants were instructed that they would see pictures on the screen and that they could advance through the images at their own pace by pressing the spacebar on the keyboard. They were asked to keep their heads as still as possible and their hands on the table for accurate eye tracking. Each participant saw a total set of nine slideshows. Individual slideshows were presented in a single randomized order with half of the participants receiving the original order and the other half receiving the reversed order.

Prior to the main task, participants completed a practice trial involving screenshots from a video of a penguin throwing a snowball, familiarizing them with the process of pressing the spacebar to watch an event unfold. After the practice trials, a 5-point calibration and validation procedure were completed. Following an opportunity to ask questions, the main task began.

In the adult experiments, the experimenter sat behind the participants, while in the child experiments, the experimenter sat beside them to ensure their attention remained focused and they did not look elsewhere. Furthermore, during the pilot study[1] conducted with children, we observed a challenge in their ability to reach the spacebar, unlike adults. Consequently, we modified the experiment in a way that enabled them to navigate through the images by clicking the mouse.

We recorded the latency between button presses as an index on the total amount of time spent on a given slide. This latency will be used as an implicit measurement indicating the perception of event boundaries. Studies conducted with adults and children have shown that moments perceived as event boundaries are viewed for a longer time (Hard et al., 2011; Kosie & Baldwin, 2021; Meyer et al., 2011; Zheng et al., 2020). Furthermore, this study validates the use of latency of the button presses with the total duration of fixations to the slide while it remained on the screen as a methodological novelty. This approach offers a more precise estimate of attention directed at each slide.

2.4 Coding

Descriptions were transcribed and coded by native Turkish speakers on ELAN (Lausberg & Sloetjes, 2009). We coded the number of linguistic units participants used when describing the event. Following Gerwien and von Stutterheim (2018), a linguistic unit was defined as one finite verb that referred to the motion event shown in the video. Additionally, we examined the semantic features of the descriptions that consisted of multiple verb phrases. Specifically, we coded whether or not descriptions consisting of multiple verb phrases consisted of at least two different path verbs (See Table 2.2 for a list of different path and manner clauses that children and adults used to describe events.). For example, descend + enter contains multiple path verbs (See example 3) whereas roll

[1] Pilot studies were conducted independently with adults and children prior to the formal data collection process.

+ enter contains manner and path distributed into different verbs (See example 4). Clauses or verb phrases that did not refer to the motion event were not counted.

Table 2.2 *Path and Manner Verbs Used by Different Age Groups*

4-year-olds		5-year-olds		Adults	
Path	Manner	Path	Manner	Path	Manner
climb	crawl	advance	hop	advance	jump
come	hop	come	jump	arrive	leap
descend	jump	descend	leap	climb	move (in triangles)
enter	roll	enter	move (in circles)	come	roll
exit	spin	exit	move (in triangles)	descend	slid
fall	walk	go	roll	enter	spin
go		pass	spin	exit	walk
pass		reach	walk	fall	zigzag
				go	
				pass	
				reach	

(3)	Top	in-di	ve	ev-e	gir-di
	ball	descend-PST	and	house-DAT	enter-PST
		verb			verb
		path			path

‘The ball descended and entered the house.

(4)	Top	yuvarlan-dı	ve	ev-e	gir-di
	ball	roll- PST	and	house-DAT	enter-PST
		verb			verb
		manner			path

‘The ball rolled and entered the house.’

For validation, we also coded the clauses used in the descriptions separately, following the coding system developed by Özyürek and colleagues (2008). This coding process considered both subordinate verbs and the main clauses (Özyürek et al., 2008). We employed this approach because in Turkish, manner is frequently conveyed through subordinate verbs, and we aimed to ensure that no information was lost when analyzing the use of multiple verb phrases. For instance, the sentence "Ball rolled down into the house." in English can be expressed in two ways in Turkish. Firstly, the manner verb is typically subordinated to the main verb using a connective, often -arak (See example 5). Secondly, manner and paths are each expressed in independent main clauses (See example 6).

(5)	Top	yuvarlan-arak	in-di	ve	ev-e	gir-di
	ball	roll-CONN	descend-PST	and	house-DAT	enter-PST
		sub. verb	verb			verb
		manner	path			path

‘The ball rollingly descended and entered the house.’

(6)	Top	yuvarlan-dı	in-di	ve	ev-e	gir-di
	ball	roll- PST	descend-PST	and	house-DAT	enter-PST
		verb	verb			verb
		manner	path			path

‘The ball rolled, descended, and entered the house.’

2.5 Preprocessing of the Eye Data in the Dwell Time Task

A message was sent from the Presentation software to the eye-tracker to indicate the onset and offset of each slide. A rectangular Area of Interest (AoI) was defined to cover each slide using SMI BeGaze software, to ensure participants' attention was directed to stimuli. Fixation durations to the AoI were computed by the SMI BeGaze software.

Using R script version 4.2.3 (R Core Team, 2023), we assessed whether participants experienced more than 45% trackloss throughout the entire task or if individual trials had over 50% trackloss. None of the participants exhibited more than 45% trackloss across all trials in the non-linguistic task; consequently, no participant was excluded from the analyses. We excluded trials with more than 50% trackloss (2.75% of all data, children: 2.15%, adults: 0.6%).

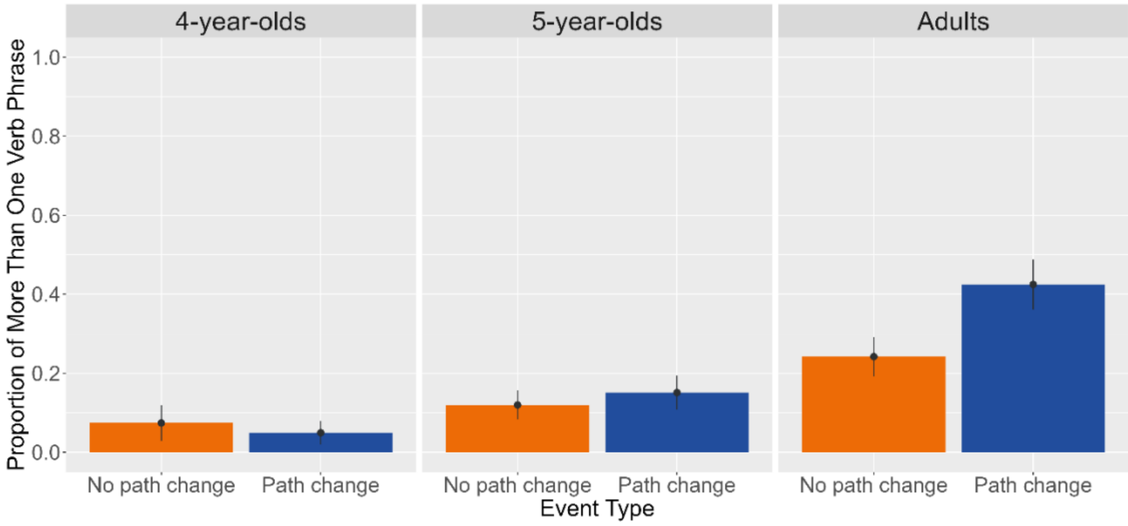
3. RESULTS

3.1 Linguistic Task

3.1.1 Analysis of More Than One Verb Phrase

The data from the linguistic task were analyzed using generalized binomial linear mixed effects models. Models were fit using the *glmer* function in the *lme4* package (version 1.1.33; Bates et al. 2015) in R (version 4.2.3; R Core Team, 2023).

Figure 3.1 *Proportion of More Than One Verb Phrase Across Event Types*



First, we examined the variation in the number of verb phrases used by participants when describing different types of events. The model tested the fixed effect of event type (path-change, no path-change) and age (adults, 4-year-olds, 5-year-olds) on the binary dependent variable of using more than one verb phrase (1 = more than one verb phrase, 0 = single verb phrase) [2] at the trial level (see Figure 3.1). The fixed effect of event type was tested with orthogonal sum-to-zero contrasts (no path change events contrast

[2] In the child data, 1% of the trials in which participants used more than one verb phrase were due to the use of multiple manner verbs.

coded as -1/2, path change events contrast coded as 1/2) (Schad et al., 2020). The fixed effect of age was tested with two orthogonal sum-to-zero contrasts. The first contrast compared 4-year-olds to 5-year-olds (4-year-olds contrast coded as -1/2, 5-year-olds contrast coded as 1/2, adults contrast coded as 0). The second contrast compared all children to adults (4-year-olds contrast coded as -1/3, 5-year-olds contrast coded as -1/3, adults contrast coded as 2/3). The model also included random intercepts for Subjects and Items (Baayen, 2008; Baayen, Davidson, & Bates, 2008).

Table 3.1 *Parameter Estimates for the Fixed Effects from the Glmer Model for Using More Than One Verb Phrase*

Effect	ß	*SE*	*z*	*p*-value
Intercept	-2.429	0.342	-7.095	<.001
Condition (No path change vs. Path change)	0.406	0.477	0.852	.394
Age (4-year-olds vs. 5-year-olds)	1.223	0.624	1.961	.050
Age (Children vs. Adults)	2.169	0.472	4.599	<.001
Condition (No path change vs. Path change): Age (4-year-olds vs. 5-year-olds)	0.303	0.916	0.331	.741
Condition (No path change vs. Path change): Age (Children vs. Adults)	1.063	0.599	1.775	.076

*Formula in R: Multiple Verb Phrase ~ Condition * Age + (1 | Subject) + (1 | Item)*

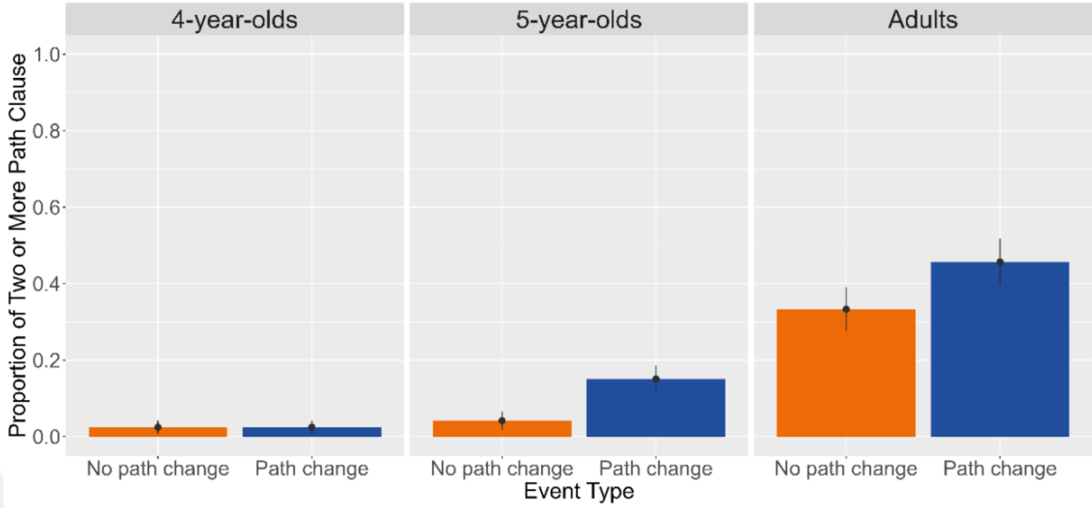
The parameter estimates for the fixed effects from the glmer model for using more than one verb phrase model are presented in Table 3.1. The fixed effect of event type was not statistically significant: proportions of more than one verb phrase descriptions were similar between events that involved a path change and events that did not involve a path change (See Figure 3.1)[3]. Furthermore, the model revealed a significant fixed effect of age in both contrasts. 5-year-olds were more likely to describe events with more than one verb phrase compared to 4-year-olds. Furthermore, adults were more likely to describe events with more than one verb phrase compared to children. No other effects or interactions were statistically significant, indicating that these patterns were similar across event type and age.

3.1.2 Analysis of Two or More Path Verbs

Next, we tested how the number of path verbs used in speech differed depending on event type and age. Prior to this analysis, we excluded 4-year-olds due to their low proportion of using two or more path verbs (M = 0.01), in contrast to 5-year-olds (M = 0.07) and adults (M = 0.29). The model tested the fixed effect of event type (path-change, no path-change) and age (adults, 4-year-olds, 5year-olds) on the binary dependent variable (1 = two or more path verbs, 0 = less than two path verbs) at the trial level (see Figure 3.2). The fixed effect of event type was tested with orthogonal sum-to-zero contrasts (no path change events contrast coded as -1/2, path change events contrast coded as 1/2) (Schad et al., 2020). The fixed effect of age was tested with orthogonal sum-to-zero contrasts (5-year-olds contrast coded as -1/2, adults contrast coded as 1/2) (Schad et al., 2020). The model also included random intercepts for Subjects only (Baayen, 2008; Baayen, Davidson, & Bates, 2008). A more complex model with random intercepts for Items resulted in a singular fit error, so this term was omitted.

[3] Due to the binary nature of the data, Figure 1 only depicts the proportion of using more than one verb phrase. The proportion of using one verb phrase or less than one verb phrase can be deduced by subtracting the proportions from 1.

Figure 3.2 *Proportion of Descriptions With at Least Two Path Verbs Across Event Types*



The parameter estimates for the fixed effects from the glmer model for using two or more path verbs model are presented in Table 3.2. Results revealed a significant fixed effect of event type: participants were more likely to use two or more path verbs in descriptions of path-change events compared to no path-change events (See Figure 3.2)[4]. Furthermore, the fixed effect of age was also significant: adults were more likely to use two or more path verbs compared to 5-year-olds. No other effects or interactions were statistically significant.

We also replicated the analysis based on the variation in the number of clauses used by participants when describing different types of events. We followed the same procedures for data exclusion, as well as model fitting as described above. See Appendix A for the parameter estimates for the fixed effects from the glmer model for using multiclause (Table 3.4) and Appendix B for using two or more path clauses (Table 3.5) models. Additionally, Appendix C presents the proportion of descriptions with

[4] Due to the binary nature of the data, Figure 3 only depicts the proportion of using two or more path verbs. The proportion of using less than two path verbs can be deduced by subtracting the proportions from 1.

multiclause (Figure 3.3) and Appendix D presents the proportion of descriptions with at least two path clauses (Figure 3.4) across event types.

Table 3.2 *Parameter Estimates for the Fixed Effects from the Glmer Model for Using Two or More Path Verbs*

Effect	ß	*SE*	*z*	*p*-value
Intercept	-2.198	0.287	-7.654	<.001
Condition (No path change vs. Path change)	1.235	0.401	3.080	.002
Age (5-year-olds vs. Adults)	2.066	0.501	4.119	<.001
Condition (No path change vs. Path change): Age (5-year-olds vs. Adults)	-0.007	0.795	-0.008	.993

*Formula in R: Two or More Path ~ Condition * Age + (1 | Subject)*

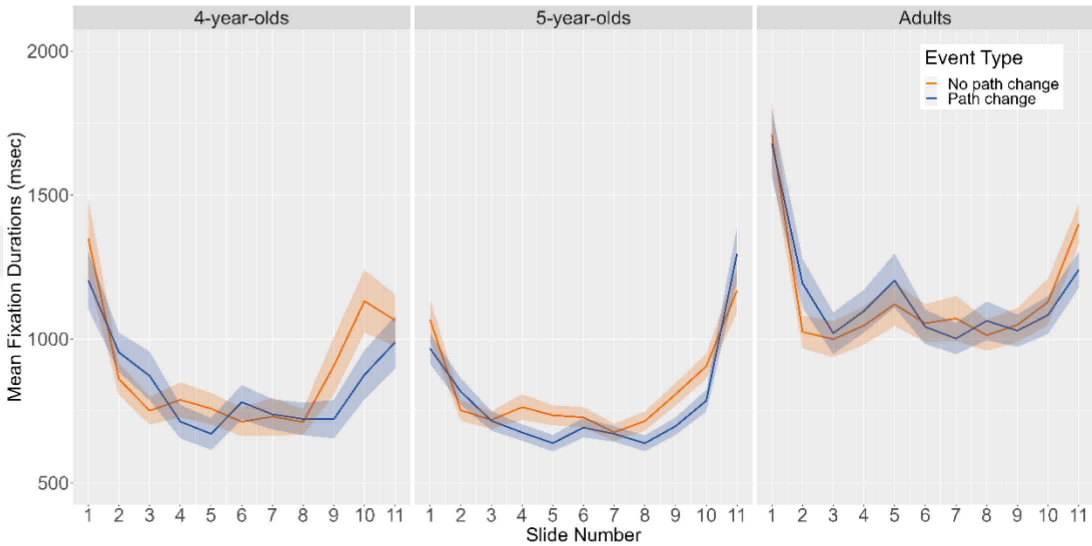
3.2 Non-linguistic Dwell Time Task

First, we tested the correlation between the two Dwell Time measures: total duration of fixation during a slide and the latency of the button presses. There was indeed a strong positive correlation between these two measures (r = .85).

We began by analyzing fixation durations as they provide a more precise estimate of attention directed at each slide. Following the previous Dwell Time studies (e.g., Hard et al., 2011; Kosie & Baldwin, 2019a, 2019b, 2021) data from the first and last slides from each slideshow were removed prior to analyses. To prepare the Dwell Time data for analyses, we adhered to standard procedures including log transformation and excluding trials that had fixation durations 3 standard deviations above or below the mean (0.66% of all data, children: 0.56%, adults: 0.1%).

Fixation durations were analyzed with a linear mixed effects model. The model was fit using the *lmer* function in the *lme4* package (version 1.1.33; Bates et al. 2015) in R (version 4.2.3; R Core Team, 2023).

Figure 3.3 *Mean Fixation Durations Across Event Types and Slides*



The model tested the fixed effect of event type (path-change, no path-change) and age (adults, 4-year-olds, 5-year-olds) on the fixation durations to slides depicting a path change and leading up to the path change as the dependent measure. The fixed effect of event type was tested with orthogonal sum-to-zero contrasts (no path change events contrast coded as -1/2, path change events contrast coded as 1/2) (Schad et al., 2020). The fixed effect of age was tested with two orthogonal sum-to-zero contrasts (Schad et al., 2020). The first contrast compared 4-year-olds to 5-year-olds (4-year-olds contrast coded as -1/2, 5-year-olds contrast coded as 1/2, adults contrast coded as 0). The second contrast compared all children to adults (4-year-olds contrast coded as -1/3, 5-year-olds contrast coded as -1/3, adults contrast coded as 2/3). The model also included random intercepts for Subjects and Items.

The parameter estimates for the fixed effects from the lmer model for the mean fixation durations model are presented in Table 3.3. The fixed effect of event type was not statistically significant: fixation durations were similar between events that involved a path change and events that did not involve a path change (see Figure 3.5). However, the fixed effect of age in the second contrast was statistically significant: fixation durations of adult participants were longer than those of child participants. No other effects or interactions were statistically significant, indicating that these patterns were similar across event types and age.

Table 3.3 *Parameter Estimates for the Fixed Effects from the Lmer Model for the Mean Fixation Durations*

Effect	ß	*SE*	*df*	*t*	*p*-value
Intercept	2.854	0.025	18.560	115.597	<.001
Condition (No path change vs. Path change)	-0.017	0.036	5.818	-0.475	.652
Age (4-year-olds vs. 5-year-olds)	-0.013	0.044	80.892	-0.286	.776
Age (Children vs. Adults)	0.182	0.037	82.424	4-903	<.001
Condition (No path change vs. Path change): Age (4-year-olds vs. 5-year-olds)	-0.027	0.030	831.734	-0.915	.360
Condition (No path change vs. Path change): Age (Children vs. Adults)	0.026	0.024	834.490	1.061	.289

*Formula in R: Mean Fixation Durations ~ Condition * Age + (1 | Subject) + (1 | Item)*

To maintain similarity to previous work (Hard et al., 2011; Kosie & Baldwin, 2021, Meyer et al., 2011; Zheng et al., 2020), we also replicated our analysis using the latency of the button presses as an index of the Dwell Time. We followed the same procedures for data transformation and exclusion, as well as model fitting in the analysis of the Dwell Time data. See Appendix E for the parameter estimates for the fixed effects from the lmer model for the mean latency of the button presses model (Table 3.6). This model also did not reveal a significant fixed effect of event type (β = 22.960, *SE* = 95.295, t = 0.241, p = .818): the latency of the button presses was similar when the slides depicted a path change as opposed to when they did not depict a path change. The fixed effect of age in the second contrast was statistically significant (β = 528.251, *SE* = 96.852, t = 5.454, p < .001): the overall latency of the button presses of adult participants was longer than child participants. No other effects or interactions were statistically significant, indicating that these patterns were similar across event types and age.

4. DISCUSSION

People readily segment continuous actions that unfold around them into event units and communicate about these events. However, the precise nature and the developmental continuity of the connection between event units in language and cognition remains debated. Some argue that semantic distinctions in language create stable differences in how speakers of different languages reason about events even when they are not using language (Levinson, 2003; Majid et al., 2004). On an alternative view, event categories are shared between speakers of different languages to a large extent (Gleitman & Papafragou, 2016; Landau et al., 2010). In the present study, we revisit this issue by testing how children and adults who speak Turkish—a verb-framed language that segments motion paths into different verb phrases—describe and perceive motion events. We asked two research questions. First, we asked to what extent linguistic differences in event segmentation influence non-linguistic event segmentation in cognition. Second, we asked whether linguistic influences on event segmentation change developmentally throughout language acquisition.

4.1 Event Segmentation in Language and Cognition

Regarding our first question, we investigated whether participants would be more likely to express events involving a path change by using multiple verb phrases compared to those that do not involve a path change. Contrary to our expectation, our Turkish-speaking participants did not significantly use more than one verb phrase when describing motion events involving a change in direction of motion compared to the motion events that did not involve a change in direction. Nevertheless, a closer inspection of participants' descriptions showed that the descriptions consisting of multiple verb phrases indeed had two or more path verbs. Furthermore, such uses were more frequent in the descriptions of events with a path change compared to events without a path change. This aligns with our predictions and confirms that our Turkish-speaking participants followed the patterns documented in previous studies. These findings establish that the segmentation of motion paths in language is a good test case for investigating the extent to which language might affect the segmentation of motion paths in non-linguistic cognition.

Next, we investigated whether these linguistic patterns would be reflected in how motion events are segmented. Recall that there were two main possibilities regarding linguistic influences on non-linguistic event segmentation. According to one possibility, Turkish speakers were expected to segment events involving a path change differently compared to those that did not. This difference would be reflected as an increase in Dwell Time on slides corresponding to the path change (for events that have one). According to an alternative possibility, the number of event units in non-linguistic cognition would not necessarily parallel the number of event units in the language. This possibility would unfold in two different ways: either participants segment events involving a path change and those that do not into an equal number of units, or they segment events into finer units in cognition before segmenting them into finer units in language. In line with the second possibility, our findings revealed that Dwell Times for slides corresponding to path changes for events that involve a path change were not significantly longer than temporally similar slides in events that did not involve a path change, and hence did not depict a change in direction of motion. This result provides evidence against the position that event units in language have stable influences on how people perceive and reason about events.

Overall, our findings seem to diverge from those of Gerwien and von Stutterheim (2018), who demonstrated that speakers of French and German differ in how they segment motion paths in cognition in line with cross-linguistic differences in motion event encoding in language. In our study, although there are linguistic differences between children and adults, there were no significant differences in the non-linguistic segmentation task. We believe these differences can be attributed to the nature of the non-linguistic event segmentation tasks used across the two studies. While the previous study used the Newtson task (1973) that measures explicit judgments of event boundaries, our study relied on the relatively more implicit Dwell Time paradigm (Hard et al., 2011). As discussed earlier, the explicit and verbal nature of the Newtson task might have encouraged participants to rely on language during event segmentation judgments, resulting in discontinuities across speakers of French and German. By contrast, our findings suggest that event units in language do not shape how coarsely or finely event units are segmented in cognition, at least under conditions that do not encourage reliance on language. Furthermore, our norming data also corroborates this explanation. Despite the fact that Dwell Times did not increase on slides corresponding to path changes, in the norming task participants were more likely to indicate the presence of a boundary for such slides compared to temporally similar slides for events that did not involve a path change. This norming data was elicited with similar (verbal and explicit) instructions as the Newtson task (1973).

Our findings align with a very recent cross-linguistic study by Lee and colleagues (2024) that investigated the relationship between conceptual representations and linguistic encoding of motion events. This study found that even though English and Turkish speakers showed clear differences in how they formed linguistic event units, these differences did not carry over to their non-linguistic unit formation. This was the case for both the Newtson task and an event individuation task.

Our findings are consistent with the findings of another cross-linguistic study on event segmentation comparing speakers of Dutch and Avatime (Defina, 2016). Unlike Dutch (a satellite-framed language), Avatime, a Kwa language spoken in Ghana, is known for its pervasive use of serial verb constructions, where two or more verbs appear consecutively, forming a compound structure that refers to a single event. However, despite this cross-linguistic difference, speakers of Avatime and Dutch did not differ in a Dwell Time segmentation task, with only an effect related to familiarity with events being observed. In another study, differences in non-linguistic event segmentation emerged when Avatime speakers were primed with serial verb constructions compared to coordinate clauses prior to the Dwell Time task. Together, these findings suggest that linguistic influences on event segmentation—if any—seem to be limited to situations in which language is explicitly or implicitly used during non-linguistic event segmentation, proposing a correspondence between how events are mentally organized during the process of formulating speech (thinking-for-speaking) and the linguistic structures used to express these events (Slobin, 1996).

4.2 Developmental Patterns in Event Segmentation

Regarding our second question, we compared 4-year-olds, 5-year-olds, and adults in both linguistic and non-linguistic tasks to explore whether there is a continuity of interactions between language on cognition in event segmentation. In the linguistic task, based on the linguistic properties of Turkish, we predicted an increase in the production of descriptions with multiple verb phrases as age increases. Specifically, we predicted that 5-year-olds would be more likely to describe events with multiple verb phrases than

4-year-olds, while adults would be more likely to describe events with multiple verb phrases than children. In line with our prediction, there was indeed a developmental trend in our results: 5-year-olds were more likely to describe events with more than one verb phrase compared to 4-year-olds. Similarly, adults were more likely to describe events with more than one verb phrase compared to children. A closer inspection of participants' descriptions showed that adults were more likely to use two or more path verbs when describing events compared to 5-year-olds. 4-year-olds rarely used two or more path verbs to describe events. These results again show a developmental pattern in the use of language-specific structures. The significant increase with age in the use of two or more path verbs in descriptions of path-change events points to a refined sensitivity to the linguistic demands of different event types as language development progresses. This sensitivity enables more precise and detailed event descriptions.

Recall that there were two main possibilities regarding the developmental continuity of linguistic influences on non-linguistic event segmentation. According to one possibility, adults and children were expected to segment events involving a path change differently compared to those that did not. Thus, the increase in Dwell Time on slides corresponding to path change (for events that have one) was expected to be higher in adults and older children compared to younger children. According to an alternative possibility, both adults and children segment events involving a path change and those that do not into an equal number of units, or that children initially segment events involving a path change into finer cognitive units before doing so linguistically. In line with the second possibility, our findings revealed that adults and children segmented events into an equal number of units. This result provides evidence for the position that event categories are shared between adults and children regardless of language acquisition. However, we found an age difference: regardless of the event type, overall fixation durations of adult participants were significantly longer than those of children. The longer fixation durations observed in adults compared to children, hint at developmental theories positing an increase in cognitive and attentional capacities with age as reflected in eye movements (see for a review: Amso & Scerif, 2015). Eyetracking paradigms, which measure the attentional state by tracking eye movements, have proven useful in studying language development in children (Tanenhaus & Trueswell, 2006; Trueswell, 2008). It is also well suited for studies with young children, allowing researchers to track their attention and visual focus as they engage in tasks (Tanenhaus & Trueswell, 2006; Trueswell, 2008). In the developmental context, studying how children's eye movements change over time can offer valuable information about the development of attentional control and cognitive processes. For example, researchers can track how children disengage from one stimulus and move to another.

4.3 Theoretical Implications

The current study investigated how linguistic patterns in event segmentation influence non-linguistic event segmentation in cognition and whether these linguistic influences developmentally change throughout language acquisition. To answer these questions, the study gathered theories of cognitive science, linguistics, and developmental psychology by examining event segmentation with linguistic and non-linguistic measures and with preschoolers and adults to understand the nature of language-cognition interactions in event segmentation. Overall, our findings provide evidence for the universalist view of the interaction between language and cognition in the following ways (Gleitman & Papafragou, 2016; Landau et al., 2010). First, the lack of parallelism between linguistic and cognitive event units suggests that cognitive event segmentation operates independently of linguistic input. Second, while developmental differences were found in how motion event components were linguistically encoded, no significant differences between the age groups were observed in the non-linguistic task in events involving a change of path versus those that did not.

Our study has implications for the acquisition of language-specific structures. Although previous studies have shown that children typically start describing events in a language-general manner and progress to mapping different linguistic elements onto different components in accordance with language-specific patterns around the age of three (Allen et al., 2007; Hickmann et al., 2009; Maguire et al., 2010; Özçalışkan & Slobin, 1999; Özyürek et al., 2008; Papafragou, Massey & Gleitman, 2002; 2006; Papafragou & Selimis, 2010), our child participants did not show an adult-like encoding in their descriptions at the age of 5. This aligns with Bunger and colleagues (2012), who found that 4-year-olds might often omit path and manner components in their descriptions. Consistent with this, our younger participants were less likely to use multiple verb phrases compared to adults. Nonetheless, a developmental trend is evident, as 5-year-olds used multiple verb phrases significantly more often than 4-year-olds, who rarely used two or more path verbs in their descriptions. The developmental trajectories observed in our study underscore the complex nature of language acquisition. As individuals mature, they acquire the linguistic tools to describe the world more richly and accurately.

Our findings converge with another line of work on motion event expressions in speech and gestures that approach the relation between verbal and non-verbal event segmentation from a different perspective. This line of work has shown that when people describe motion events, each co-speech gesture expresses the semantic information (i.e., path, manner, or both) encoded within a verbal clause in the accompanying speech (Kita & Özyürek, 2003). For example, speakers of Turkish (and other verb-framed languages) use separate gestures to express manner and path. However, when people are asked to describe motion events only using gestures without accompanying speech, speakers of both Turkish and English conflated path and manner into a single gesture (Özçalışkan et al., 2016; 2018). This contrast in the gestures produced with and without accompanying speech is also reflected in our linguistic and non-linguistic tasks: our Turkish-speaking participants perceived multiple motion paths as part of a single event unit even though they distributed this information into multiple units in language. Overall, our results strongly suggest that Turkish speakers can flexibly shift between different units of representations when segmenting events in language and cognition.

Several aspects of our findings require further consideration. First, Event Segmentation Theory, as detailed in the Introduction section, suggests that an event boundary is triggered when there are changes in core features of an event, such as direction/orientation, and what will happen next is no longer predictable. However, in our study, the fixation durations to slides corresponding to path changes were not significantly higher compared to temporally similar slides without path changes. This observation suggests that the predictability of path change moments might be higher than anticipated. It raises questions about whether the participants perceived these changes as significant enough to constitute an event boundary.

Another important aspect to consider is the significant yet relatively low boundary agreements in our norming data, with 46% for path-change events and 30% for events without path changes. This was also reflected in our findings of the linguistic task: adult participants' mean proportion of expressing path-change events with two or more path verbs was slightly lower than 50%. This may be due to our stimuli in the following ways. We tried to minimize the video durations, and therefore the number of slides in order to provide particularly child-friendly stimuli. This may have been a disadvantage for the Dwell Time paradigm. Additionally, the path changes in our events might not have been sufficiently salient.

4.4 Methodological and Practical Implications

To our knowledge, this is the first study that tested event segmentation using both linguistic and non-linguistic measurements across preschoolers and adults. In this study, we integrated methodologies from linguistics, cognitive science, and developmental psychology. In the linguistic task, we analyzed event descriptions with the help of influential linguistic theories (Talmy, 1985; 2000). In the non-linguistic task, we used the Dwell Time paradigm (Hard et al., 2011) which offers an implicit measurement for event segmentation. Furthermore, to see the developmental patterns, we tested participants from various age groups, including preschoolers and adults. This approach allowed us to examine developmental patterns across different stages of language acquisition, highlighting findings from understudied preschoolers and providing a bridge between studies on infants and adults. Moreover, focusing on Turkish, a verb-framed language that is less represented in existing research, enhances our understanding of the language-cognition interface across different linguistic contexts.

This study introduced two methodological novelties. First, we validated dwell times with fixation durations. The high correlation between the two measures shows the usefulness of fixation durations as a reliable index of attention allocation to individual slides. This is particularly relevant for child participants; whose attentional capacities naturally differ from adults. Second, we deepened our investigation of multiple verb phrase descriptions by analyzing the content of these phrases regarding whether path or manner phrases were produced. This allowed us to determine whether path changes are encoded in the descriptions, or if participants simply used multiple verb phrases with path and manner distributed across different verb phrases.

The current study offers a few practical implications. Event segmentation relates to other cognitive skills such as predictive processing, imitation, event categorization, and word learning (Stahl et al., 2014). For instance, segmenting events may allow children to be able to predict upcoming events by forming hierarchical event representations in a meaningful context, like a bedtime routine that includes bathing, brushing teeth, and reading (Stahl et al., 2014). This skill might also boost their memory of these events (Zacks & Swallow, 2007). Additionally, recognizing and understanding the sequence of events is crucial for children when they begin to enact these actions, such as bathing by segmenting actions into units like washing and then rinsing. Thus, predicting the action sequences could be key to how children learn to both comprehend and perform actions (Stahl et al., 2014). Considering these practical benefits of event segmentation, it would be beneficial to incorporate strategies that align with children's developing cognitive and linguistic capabilities in caregiving or educational settings.

4.5 Limitations and Future Directions

The current study has some limitations that should be considered by future research. As noted in our theoretical implications, the relatively low agreement on boundaries may stem from the characteristics of our stimuli. While longer stimuli could potentially enhance the effectiveness of the Dwell Time paradigm, it remains important to use child-friendly stimuli. Moreover, future studies could use more salient examples of path changes, such as entering and exiting a house or climbing up and then descending a hill.

Although our study provides novel insights into the interaction between language and cognition, several issues remain open for further exploration. While this study focuses on Turkish speakers, its methodology and findings could serve as a basis for cross-linguistic studies. Such comparisons could further elucidate the interaction between linguistic and cognitive processes across different linguistic communities. Additionally, employing a longitudinal design could offer valuable insights into how event segmentation patterns individually change over time. Moreover, conducting this study with bilingual children, especially those fluent in both verb-framed and satellite-framed languages, would deepen our understanding of the interaction between different linguistic processes. Our findings also suggest the possibility of conducting similar studies with younger and/or older children, as our youngest participants did not show significant differences in the use of multiple verb phrases across different event types.

Moreover, most empirical data is derived from studies of motion events. The relationship between event segmentation in language and cognition requires further investigation beyond motion events to determine if these findings hold across different event types. Furthermore, although our stimuli consisted of animated videos, incorporating stimuli that depict realistic events could enhance the ecological validity of the findings, as event segmentation naturally occurs in our daily lives. Finally, it's crucial to acknowledge that communication about events is typically multimodal (Kita & Özyürek, 2003). It might be useful to study the gestures along with the speech to see whether they mirror the event components in language. For example, the study by Özyürek and colleagues (2008) found that children learning English and Turkish showed differences in their speech patterns from age 3 onward, with English speakers using one clause and Turkish speakers using two clauses to express manner and path components. However, the gestures of English- and Turkish-speaking children looked similar at ages 3 and 5, with both groups using separate gestures for manner and path. Their gestures diverged at age 9 and in adulthood. This suggests a gradual shift towards language-specificity during speaking, indicating that gestures provide insights into the development of language-specific event representations. However, in our study, we could not video-recorded children's event descriptions due to practical reasons.

4.6 Conclusion

The current study investigated how Turkish-speaking 4- and 5-year-olds and adults segmented motion events in language and cognition. Our findings revealed that participants described motion events that involve a change in direction with two or more path verbs more frequently than events that do not involve a change in direction. Notably, a developmental trend emerged in the linguistic descriptions: adults and older children used multiple verb phrases more frequently than younger children. This trend was particularly evident in the use of two or more path verbs. Adults were more likely to describe events with two or more path verbs compared to 5-year-olds. 4-year-olds rarely used multiple path verbs, indicating they have not yet acquired encoding different motion paths into different verb phrases. However, in a non-linguistic segmentation task, their Dwell Times did not differ depending on whether the events involved a change in direction. This suggests that children and adults segment the events similarly, but adults have a richer linguistic capacity to describe these segments, particularly when it comes to path changes. By investigating the complex relationship between linguistic expressions and cognitive segmentation across different age groups, our study underscores the dynamic nature of event cognition and its developmental trajectory.

REFERENCES

Aktan-Erciyes, A., & Göksun, T. (2019). Early event understanding predicts later verb comprehension and motion event lexicalization. *Developmental Psychology*, *55* (11), 2249. https://psycnet.apa.org/doi/10.1037/dev0000804

Allen, S., Özyürek, A., Kita, S., Brown, A., Furman, R., Ishizuka, T., & Fujii, M. (2007). Language-specific and universal influences in children's syntactic packaging of manner and path: A comparison of English, Japanese, and Turkish. *Cognition*, *102*(1), 16–48. https://doi.org/10.1016/j.cognition.2005.12.006

Amso, D., & Scerif, G. (2015). The attentive brain: insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, *16*(10), 606-619. https://doi.org/10.1038/nrn4025

Baayen, R. H. (2008). Analyzing linguistic data a practical introduction to statistics using R. New York, NY: Cambridge University Press.

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. https://doi.org/10.1016/j.jml.2007.12.005

Baldwin, D. A., Baird, J. A., Saylor, M. M., & Clark, M. A. (2001). Infants parse dynamic action. *Child Development*, *72*(3), 708–717. https://doi.org/10.1111/1467-8624.00310

Bates, D., Machler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Bohnemeyer, J., Enfield, N. J., Essegbey, J., Ibarretxe-Antunano, Iraide., Kita, S., Lüpke, Friederike., & Ameka, F. K. (2007). Principles of event segmentation in language: The case of motion events. *Language*, *83*(3), 495–532. https://doi.org/10.1353/lan.2007.0116

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, *122*(2), 135–149. https://doi.org/10.1016/j.cognition.2011.10.002

Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2021). How children attend to events before speaking: Crosslinguistic evidence from the motion domain. *Glossa: A Journal of General Linguistics*, *6*(1), Article 1. https://doi.org/10.5334/gjgl.1210

Cardini, F. E. (2010). Evidence against Whorfian effects in motion conceptualization. *Journal of Pragmatics*, *42* (5), 1442-1459. https://doi.org/10.1016/j.pragma.2009.09.017

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *41*(1-3), 83-121. https://doi.org/10.1016/0010-0277(91)90033-Z

Cohen, L. B., & Oakes, L. M. (1993). How infants perceive a simple causal event. *Developmental Psychology*, *29*(3), 421-433. https://doi.org/10.1037/0012-1649.29.3.421

Defina, R. (2016) Events in Language and Thought: The Case of Serial Verb Constructions in Avatime, Radboud University Nijmegen (PhD Thesis)

Feinmann, D. (2020). Language and thought in the motion domain: Methodological considerations and new empirical evidence. *Journal of Psycholinguistic Research*, *49*(1), 1-29. https://doi.org/10.1007/s10936-019-09668-5

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, *83*(1), 49–79. https://doi.org/10.1016/S0010-0277(01)00166-4

Gerwien, J., & von Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, *180*, 225–237. https://doi.org/10.1016/j.cognition.2018.07.008

Gleitman, L., & Papafragou, A. (2012). New perspectives on language and thought. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp.543–568). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199734689.013.0028

Göksun, T., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Konishi, H., & Okada, H. (2011). Who is crossing where? Infants' discrimination of figures and grounds in events. *Cognition*, *121*(2),176-195. https://doi.org/10.1016/j.cognition.2011.07.002

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of Experimental Psychology: General*, *140*(4), 586–604. https://doi.org/10.1037/a0024310

Hickmann, M., Taranne, P., & Bonnet, P. (2009). Motion in first language acquisition: Manner and Path in French and English child language. *Journal of Child Language*, *36*(4), 705-741. https://doi.org/10.1017/S0305000908009215

Kita, S., & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*(1), 16–32. https://doi.org/10.1016/S0749-596X(02)00505-3

Kosie, J. E., & Baldwin, D. (2019a). Attentional profiles linked to event segmentation are robust to missing information. *Cognitive Research: Principles and Implications*, *4*(1), 1-18. https://doi.org/10.1186/s41235-019-0157-4

Kosie, J. E., & Baldwin, D. (2019b). Attention rapidly reorganizes to naturally occurring structure in a novel activity sequence. *Cognition*, *182*, 31-44. https://doi.org/10.1016/j.cognition.2018.09.004

Kosie, J. E., & Baldwin, D. A. (2021). Dwell Times showcase how goal structure informs preschoolers' analysis of unfolding motion patterns. *Child Development*, *92*(6), 2235–2243. https://doi.org/10.1111/cdev.13661

Landau, B., Dessalegn, B., & Goldberg, A. M. (2010). Language and space: Momentary interactions. In P. Chilton & V. Evans (Eds.), *Language, cognition, and space: The state of the art and new directions* (pp. 51–78). London: Equinox Publishing.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods, 41*(3), 841–849. https://doi.org/10.3758/BRM.41.3.841

Lee, S., Ünal, E., & Papafragou, A. (2024). Forming Event Units in Language and Cognition: A Cross-linguistic Investigation. Proceedings of the Annual Meeting of the Cognitive Science Society, 46. Retrieved from https://escholarship.org/uc/item/8wq230f5

Levelt, W. (1989). *Speaking*. MIT Press.

Levinson, S. C. (2003). *Space in language and cognition explorations in cognitive diversity*. Cambridge University Press.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., Imai, M., Haryu, E., Vanegas, S., ... & Sanchez-Davis, B. (2010). A developmental shift from similar to language-specific strategies in verb acquisition: A comparison of English, Spanish, and Japanese. *Cognition*, *114*(3), 299-319. https://doi.org/10.1016/j.cognition.2009.10.002

Majid, A., Bowerman, M., Kita, S., Haun, D. B., & Levinson, S. C. (2004). Can language restructure cognition? The case for space. *Trends in Cognitive Sciences*, *8*(3), 108-114. https://doi.org/10.1016/j.tics.2004.01.003

Meyer, M., Baldwin, D. A., & Sage, K. (2011). Assessing young children's hierarchical action segmentation. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33, No. 33).

Naigles, L. R., Eisenberg, A. R., Kako, E. T., Highter, M., & McGraw, N. (1998). Speaking of motion: Verb use in English and Spanish. *Language and Cognitive Processes*, *13*(5), 521-549. https://doi.org/10.1080/016909698386429

Newtson, D. (1973). Attribution and the unit of perception of ongoing behavior. *Journal of Personality and Social Psychology*, *28*(1), 28–38. https://doi.org/10.1037/h0035584

Newtson, D., & Engquist, G. (1976). The perceptual organization of ongoing behavior. *Journal of Experimental Social Psychology*, *12* (5), 436-450. https://doi.org/10.1016/0022-1031(76)90076-7

Özçalışkan, Ş. (2015). Ways of crossing a spatial boundary in typologically distinct languages. *Applied Psycholinguistics*, *36*(2), 485-508.

https://doi.org/10.1017/S0142716413000325

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2016). Does language shape silent gesture? *Cognition*, *148*, 10–18. https://doi.org/10.1016/j.cognition.2015.12.001

Özçalışkan, Ş., Lucero, C., & Goldin-Meadow, S. (2018). Blind speakers show language-specific patterns in co-speech gesture but not silent gesture. *Cognitive Science, 42(3),* 1001–1014. https://doi.org/10.1111/cogs.12502

Özçalışkan, Ş., & Slobin, D. I. (1999). Learning how to search for the frog: Expression of manner of motion in English, Spanish, and Turkish. In *Proceedings of the 23rd annual Boston University conference on language development* (Vol. 2, pp. 541-552). Somerville, MA: Cascadilla Press.

Özçalışkan, Ş., & Slobin, D. I. (2003). Codability effects on the expression of manner of motion in English and Turkish. In A. S. Özsoy, M. Nakipoglu-Demiralp, E. Erguvanli-Taylan & A. Aksu-Koç (Eds.), *Studies in Turkish Linguistics* (pp. 259–270). Bogazici University Press.

Özyürek, A., Kita, S., Allen, S., Brown, A., Furman, R., & Ishizuka, T. (2008). Development of cross-linguistic variation in speech and gesture: Motion events in English and Turkish. *Developmental Psychology*, *44*(4), 1040–1054. https://doi.org/10.1037/0012-1649.44.4.1040

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*(1), 155–184. https://doi.org/10.1016/j.cognition.2008.02.007

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle,'n'roll: The representation of motion in language and cognition. *Cognition*, *84*(2), 189–219. https://doi.org/10.1016/S0010-0277(02)00046-X

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition, 98*(3), B75–B87. https://doi.org/10.1016/j.cognition. 2005.05.005

Papafragou, A., & Selimis, S. (2010). Event categorization and language: A cross-linguistic study of motion. *Language and Cognitive Processes, 25*(2), 224-260. https://doi.org/10.1080/01690960903017000

Pruden, S.M., Hirsh-Pasek, K., Maguire, M.J., & Meyer, M.A. (2004). Foundations of verb learning: Infants form categories of path and manner in motion events. In A. Brugos, L. Micciulla, & C.E. Smith (Eds.), *Proceedings of the 28th annual Boston University Conference on Language Development* (pp.461-472). Somerville, MA: Cascadilla Press.

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. S. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. *Cognition, 108*(3), 825-830. https://doi.org/10.1016/j.cognition.2008.04.009

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, *17*, 133-140. https://doi.org/10.1016%2Fj.cobeha.2017.08.006

Sage, K. D., & Baldwin, D. (2014). Looking to the hands: Where we dwell in complex manual sequences. *Visual Cognition*, *22*(8), 1092-1104. https://doi.org/10.1080/13506285.2014.962123

Sargent, J. Q., Zacks, J. M., Hambrick, D. Z., Zacks, R. T., Kurby, C. A., Bailey, H. R., ... & Beck, T. M. (2013). Event segmentation ability uniquely predicts event memory. *Cognition*, *129* (2), 241-255. https://doi.org/10.1016/j.cognition.2013.07.002

Sasmita, K., & Swallow, K. M. (2023). Measuring event segmentation: An investigation into the stability of event boundary agreement across groups. *Behavior Research Methods*, *55*(1), 428-447. https://doi.org/10.3758/s13428-022-01832-5

Saylor, M. M., Baldwin, D. A., Baird, J. A., & LaBounty, J. (2007). Infants' online segmentation of dynamic human action. *Journal of Cognition and Development, 8*(1), 113–128. https://doi.org/10.1207/s15327647jcd0801_6

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Slobin, D. I. (1996). From “thought and language” to “thinking for speaking.” In J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 57–149). Cambridge University Press.

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development*, *85*(5), 1821-1826. https://doi.org/10.1111/cdev.12247

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Grammatical categories and the lexicon* (pp. 57–149). Cambridge University Press.

Talmy, L. (2000). A typology of event integration. In *Toward a cognitive semantics: Typology and process in concept structuring* (Vol. 2, pp. 213–288). MIT Press.

Tanenhaus, M. K., & Trueswell, J. C. (2006). Eye movements and spoken language comprehension. In *Handbook of Psycholinguistics* (pp. 863-900). Academic Press.

Trueswell, J. (2008). Using eye movements as a developmental measure within psycholinguistics. In I. A. Sekerina, E. M. Fernández & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing.* (pp. 73-96). Amsterdam: John Benjamins.

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, *63*(1), 64–82. https://doi.org/10.1016/j.jml.2010.02.006

Ünal, E., Mamus, E., Özyürek, A. (2023). Multimodal encoding of motion events in speech, gesture and cognition. *Language and Cognition*, 1-20. https://doi.org/10.1017/langcog.2023.61

Winaver, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, *104*(19), 7780-7785. https://doi.org/10.1073/pnas.0701644104

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253-265. https://doi.org/10.1002/wcs.104

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science*, *28*(6), 979-1008. https://doi.org/10.1207/s15516709cog2806_5

Zacks, J. M., Kurby, C. A., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction error associated with the perceptual segmentation of naturalistic events. Journal of

Cognitive Neuroscience, *23*(12), 4057-4066. https://doi.org/10.1162/jocn_a_00078

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, *16*(2), 80-84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General*, *130*(1), 29-58. https://psycnet.apa.org/doi/10.1037/0096-3445.130.1.29

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, *133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General*, *138*(2), 307–327. https://doi.org/10.1037/a0015305

Zheng, Y., Zacks, J. M., & Markson, L. (2020). The development of event perception and memory. *Cognitive Development*, *54*, 100848. https://doi.org/10.1016/j.cogdev.2020.100848

APPENDICES

APPENDIX A. FIXED EFFECT ESTIMATES FOR GLMER MODEL OF MULTICLAUSE USE

Table 3.4 presents the parameter estimates for the fixed effects from the glmer model used to analyze the use of multiclause structures. This analysis, along with the subsequent one, serves as validation for the analysis of verb phrases.

Table 3.4 *Parameter Estimates for the Fixed Effects from the Glmer Model for Using Multiclause*

Effect	ß	*SE*	*z*	*p*-value
Intercept	0.679	0.353	1.924	.054
Condition (No path change vs. Path change)	0.468	0.602	0.778	.437
Age (4-year-olds vs. 5-year-olds)	1.187	0.533	2.228	.026
Age (Children vs. Adults)	3.670	0.565	6.498	<.001
Condition (No path change vs. Path change): Age (4-year-olds vs. 5-year-olds)	0.926	0.573	1.616	.106
Condition (No path change vs. Path change): Age (Children vs. Adults)	0.723	0.638	1.134	.257

*Formula in R: Multiclause ~ Condition * Age + (1 | Subject) + (1 | Item)*

APPENDIX B. FIXED EFFECT ESTIMATES FOR GLMER MODEL OF MULTIPLE PATH CLAUSE USE

Table 3.5 presents the parameter estimates for the fixed effects from the glmer model used to analyze the use of two or more path clauses.

Table 3.5 *Parameter Estimates for the Fixed Effects from the Glmer Model for the Using Two or More Path Clause*

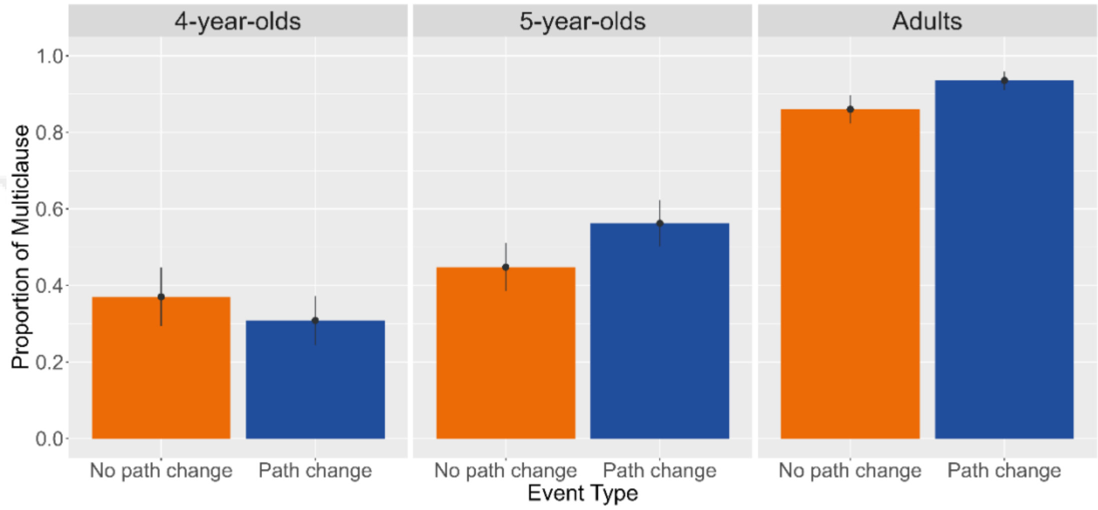
Effect	ß	*SE*	*z*	*p*-value
Intercept	-2.198	0.287	-7.654	<.001
Condition (No path change vs. Path change)	1.235	0.401	3.080	.002
Age (5-year-olds vs. Adults)	2.066	0.501	4.119	<.001
Condition (No path change vs. Path change): Age (5-year-olds vs. Adults)	-0.007	0.795	-0.008	.993

*Formula in R: Two or More Path Clause ~ Condition * Age + (1 | Subject)*

APPENDIX C. PROPORTION OF MULTICLAUSE USE

Figure 3.4 presents the proportion of multiclause across event types. This figure, along with the subsequent one, serves as validation for the analysis of verb phrases.

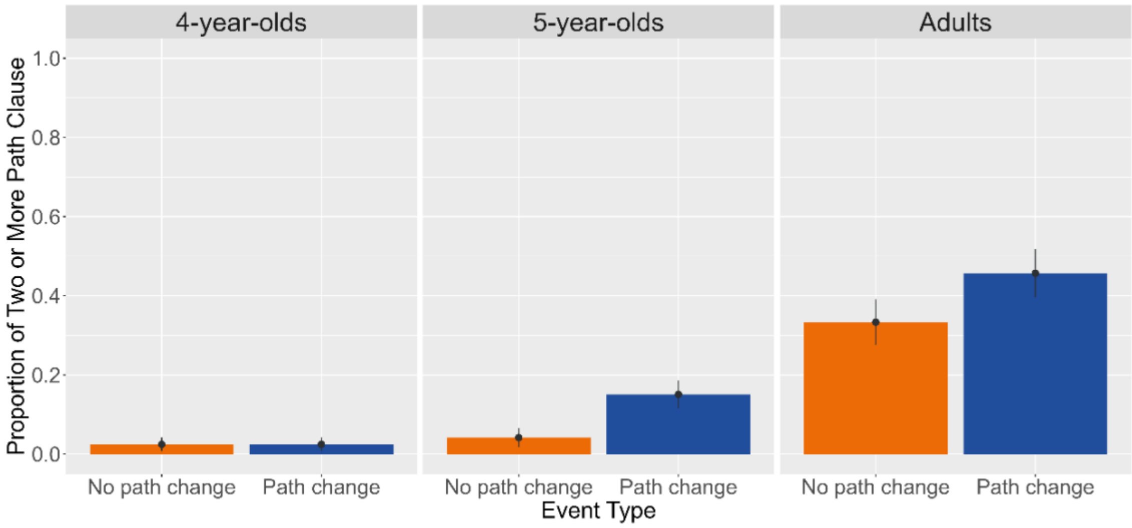
Figure 3.4 *Proportion of Multiclause Across Event Types*



APPENDIX D. PROPORTION OF MULTIPLE PATH CLAUSE USE

Figure 3.5 presents the proportion of descriptions with at least two path clauses across event types.

Figure 3.5 *Proportion of Descriptions With At Least Two Path Clauses Across Event Types*



APPENDIX E. FIXED EFFECT ESTIMATES FOR GLMER MODEL OF BUTTON PRESSES

Table 3.6 presents the parameter estimates for the fixed effects from the lmer model used to analyze the mean latency of the button presses. This analysis serves as validation for the analysis of fixation durations.

Table 3.6 *Parameter Estimates for the Fixed Effects from the Lmer Model for the Mean Latency of the Button Presses*

Effect	ß	*SE*	*df*	*t*	*p*-value
Intercept	887.621	64.341	18.059	13.796	<.001
Condition (No path change vs. Path change)	22.960	95.295	5.819	0.241	.818
Age (4-year-olds vs. 5-year-olds)	-126.667	115.367	78.216	-1.098	.276
Age (Children vs. Adults)	528.251	96.852	77.054	5.454	<.001
Condition (No path change vs. Path change): Age (4-year-olds vs. 5-year-olds)	-70.416	84.476	830.710	-0.834	.405
Condition (No path change vs. Path change): Age (Children vs. Adults)	83.626	69.123	828.651	1.210	.227

*Formula in R: Mean Latency of the Button Presses ~ Condition * Age + (1 | Subject) + (1 | Item)*