AN INVESTIGATION OF TURKISH STATIC SPATIAL SEMANTICS IN TERMS OF EXPRESSION VARIETY: AN EYE TRACKING STUDY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

ŞEYMA NUR ERTEKİN

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

AUGUST 2021

AN INVESTIGATION OF TURKISH STATIC SPATIAL SEMANTICS IN TERMS OF EXPRESSION VARIETY: AN EYE TRACKING STUDY

Submitted by ŞEYMA NUR ERTEKİN in partial fulfillment of the requirements for the degree of Master of Science in Cognitive Science Department, Middle East Technical University by,

Date: _*02/08/2021*

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

Name, Last name : **Şeyma Nur Ertekin**

Signature : ____________________

ABSTRACT

AN INVESTIGATION OF TURKISH STATIC SPATIAL SEMANTICS IN TERMS OF LEXICAL VARIETY: AN EYE TRACKING STUDY

Ertekin, Şeyma Nur

MSc., Department of Cognitive Sciences

Supervisor: Assist. Prof. Dr. Cengiz Acartürk

August 2021, 87 pages

The semantics of spatial terms has been attracting the attention of researchers for the past several decades. As an understudied language, Turkish presents an appropriate test bed for studying the generalizability of semantic characterization of spatial terms across languages. Turkish also exhibits unique characteristics, such as the use of locative case markers and being an agglutinative language. The present study reports an eye-tracking investigation of comprehension of spatial terms in Turkish by employing Topological Relations in Picture Series (Bowerman & Pederson, 1992). The major research question of the study is the presence of a relationship between the variety of spatial expressions produced by native speakers and fixation patterns on the stimuli. The findings reveal that the richness of the spatial expression is related to longer and more frequent fixations on the stimuli. The findings also show that the presence of a locative case marker in the utterances plays a role in this relation. We also investigated whether Turkish native speakers exhibit any sensitivity on *Core* and *Non-Core* distinction in spatial term semantics (Landau, 2017). Our findings showed that some of the *Non-Core* term categories reveal more variety of spatial expressions and longer fixation durations on the *Figure*, supporting cross-linguistic *Core* categorization.

Keywords: Spatial Language, Eye Tracking, Semantics

ÖZ

TÜRKÇEDE STATİK UZAMSAL DİL SEMANTİĞİNİN KELİME ÇEŞİTLİLİĞİ AÇISINDAN İNCELENMESİ: BİR GÖZ HAREKETLERİ TAKİP ÇALIŞMASI

Ertekin, Şeyma Nur
Yüksek Lisans, Bilişsel Bilimler Bölümü
Tez Yöneticisi: Doç. Dr. Cengiz Acartürk

Ağustos 2021, 87 sayfa

Uzamsal terimlerin semantiği, son zamanlarda araştırmacıların dikkatini çekmektedir. Yeterince çalışılmamış bir dil olarak Türkçe, diller arasında uzamsal terimlerin anlamsal nitelendirilmesinin genelleştirilebilirliğini incelemek için uygun bir test ortamı sunar. Türkçe aynı zamanda ismin hallerinden bulunma halinin kullanılması ve sondan eklemeli bir dil olması gibi benzersiz özellikler de sergilemektedir. Bu çalışmada, Resim Dizilerinde Topolojik İlişkiler (Topological Relations in Picture Series, Bowerman ve Pederson, 1992) uyaran serisini kullanılarak Türkçe uzamsal terimlerin anlaşılmasına yönelik bir göz izleme araştırması rapor edilmektedir. Çalışmanın temel araştırma sorusu, anadili Türkçe olan kişiler tarafından üretilen çeşitli uzamsal ifadeler ile uyaranlar üzerindeki sabitleme modelleri arasında bir ilişkinin var olup olmadığıdır. Bulgular, uzamsal ifade zenginliğinin, uyaranlara daha uzun ve daha çok fiksasyonlar ile ilintili olduğunu ortaya koymaktadır. Bulgular ayrıca, ifadelerde ismin bulunma eklerinin varlığının bu ilişkide rol oynadığını göstermektedir. Ayrıca, bu çalışmada anadili Türkçe olan bireylerin, uzamsal terim semantiğindeki *Çekirdek* 'Core' ve *Çekirdek Olmayan* 'Non-Core' ayrımına herhangi bir duyarlılık gösterip göstermediğini de araştırdık (Landau, 2017). Bulgularımız, *Çekirdek Olmayan* 'Non-Core' terim sınıflarının bazılarının daha fazla çeşitlilikte uzamsal ifadeler ve *Konumlanan* üzerinde daha uzun sabitleme sürelerine neden olduğunu ve böylece diller arası çekirdek sınıflandırmayı desteklediğini gösterdi.

Anahtar Sözcükler: Uzamsal Dil, Göz İzleme, Anlambilim

To My Family

ACKNOWLEDGEMENTS

First of all, I would like to express my deepest gratitude to my supervisor, Assoc. Prof. Dr. Cengiz Acartürk, for his guidance, encouragement, advice, and time. He inspired me so much with my thesis. I learned a lot from my supervisor, and I feel lucky to work with him.

Besides my supervisor, I would like to thank the entire Cognitive Science Department for helping me understand cognitive science deeper, providing a unique and understanding place.

I would also like to thank the people in the laboratory who showed my way around in a warm environment.

I thank my friends from the Cognitive Science department, Aysun, Gizem, and Şura, for being with me in this complicated journey, and my friends located all around the world, Aylin, Nursima, and Gülnur, for always supporting me, listening to me, and showing me that distance does not matter.

To my amazing family, Ahmet Hamdi Ertekin and Medine Ertekin, I thank them for raising me, encouraging me, supporting me, and loving me. To my brother, Mustafa Ensar Ertekin, I thank him for his support, encouragement, and patience throughout my life, especially the last few years. Lastly, but most importantly, my success is not but through Him. [11:88]

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABL	Ablative Case
DAT	Dative Case
$\mathbf{F_{avg}}$	Mean Fixation Duration on the *Figure*
$\mathbf{F_{count}}$	Fixation Count on the *Figure*
$\mathbf{F_{sum}}$	Total Fixation Duration on the *Figure*
FoR	Frame of Reference
$\mathbf{G_{avg}}$	Mean Fixation Duration on the *Ground*
$\mathbf{G_{count}}$	Fixation Count on the *Ground*
$\mathbf{G_{sum}}$	Total Fixation Duration on the *Ground*
LOC	Locative Case
POSS	Possessive
SG	Singular
$\mathbf{T_{avg}}$	Mean Fixation Duration on the *Text*
$\mathbf{T_{count}}$	Fixation Count on the *Text*
$\mathbf{T_{sum}}$	Total Fixation Duration on the *Text*
TRPS	Topological Relations in Picture Series
TSL	Turkish Sign Language

CHAPTER 1

INTRODUCTION

1.1.Aim of the Study

The aim of the present study is twofold. The first is to analyze spatial term semantics. For this, we investigated verbal descriptions and spatial term variety of static, spatial scenes in Turkish. Spatial language studies have two sub-units; dynamic spatial scene studies and static spatial scene studies (Levinson, 2003; Talmy, 2000). The present study focuses on static spatial scenes. Moreover, we focus on the analysis of the variety (richness) of spatial expressions rather than the analysis of the frequencies of spatial terms used in the verbal descriptions. The second aim of the present study is to employ the eye-tracking methodology for investigating visual information extraction processes that take place during the course of verbal descriptions.

For this, we recorded the eye movements of 36 participants during their verbal description of a visual scene battery (Topological Relations in Picture Series, Bowerman & Pederson, 1992). The battery consists of 71 line drawings depicting a *Figure* in orange and a *Ground* in black in various spatial configurations, such as *Containment*, S*upport*, and *Attachment*. The *Figure* is used to describe the object to be located, and the *Ground* is used to describe another object for location (Landau & Jackendoff, 1993), as explained below in more detail. The participants answered a prompt question, 'Where is X?', where X is the *Figure*.

In terms of its methodology, the present study employed psycholinguistic and cognitive experimentation techniques, in particular production tasks (Carlson & Hill, 2007) and eye-tracking methodology. In a production task, participants make verbal descriptions of spatial configurations of objects. We employed an open-ended task with no restrictions.[1] A major advantage of using the open-ended approach is that it is

[1] The alternative design was a fixed task, in which the participants are restricted to use terms from a limited set. We left this alternative to future work.

not certain if the participants' selection corresponds to the exact answer the participants want to give in the fixed-set approach (Carlson & Hill, 2007). We also recorded participants' eye movement patterns and response time to analyze fixation counts, fixation duration in each fixation, and overall fixation duration on a particular area on the stimuli.

We analyzed the variety of the verbal descriptions, as well as the variety of the spatial terms, the locative suffixes, the adpositions, and the reference objects. We also analyzed the *Core* vs. *Non-Core* distinction in the sub-groups of *Containment* and *Support* types of spatial terms (Landau, 2017; 2020).

1.2. Significance of the Study

Spatial language is a crucial means of expressing their daily interaction in space; thus, it is one of the important research areas of cognitive science and psycholinguistics. The differences of spatial expressions across languages make spatial language an interesting topic for understanding human cognitive processes, in particular the conceptualization of space. It has been widely studied in terms of the universality of spatial terms, childrens' spatial term acquisition, frames of references, and frequency of spatial terms for both dynamic and static spatial scenes. Recently, there exist a few studies on the investigation of Turkish from the perspective of the conceptualization of space (e.g., Atak, 2018). Most of the available studies mainly focus on dynamic spatial language (Arık, 2017; Furman, Özyürek, & Küntay, 2010; Toplu, 2011). Turkish has been investigated in cross-linguistic studies with a limited number of native speakers (Johanson & Papafragou, 2014). The major focus in the previous works on Turkish spatial terms mainly was children's language development and Turkish sign language (Johnson & Slobin, 1979; Arık, 2003; 2009; Sümer et al., 2012; 2014).

This thesis has some significance among spatial language studies. First, instead of frequencies (of daily use) of spatial expressions, the present thesis studies the diversity of utterance (i.e., the variety of spatial expressions) in each stimulus (i.e., a picture in the Topological Relations in Picture Series, TRPS, Bowerman & Pederson, 1992). Second, most of the work on Turkish is about dynamic spatial language and the verb-framed nature of Turkish. Nevertheless, the present thesis deals with static spatial language with the open-ended method, which provides more freedom to the participants, thus providing higher environmental validity in terms of its methodology. Hence, it gives more comprehensive information data on static spatial language. For example, the present study allows participants to use locative suffixes, which have not been included in Turkish spatial language studies before. Third, to our knowledge, *Core* and *Non-Core* distinction made by Landau (2017, 2020) has not been studied in Turkish. From a methodological point of view, a unique characteristic of the study is the use of eye-tracking methodology during the course of verbal descriptions of static scenes in TRPS.

1.3.Research Question of the Study

The present thesis employed eye-tracking to investigate spatial expressions of Turkish speakers in static scenes. Our main aim is to use the eye-movement methodology to study the variety of spatial expressions. The open-ended design with the eye-tracking methodology of our setting allows comprehensive verbal expressions from participants and provides pupil movements of participants within the scope of the following research questions:

1. What is the relationship between oculomotor and behavioral variables (e.g., fixation duration, fixation count, response time) and stimuli components (i.e., the *Figure* and the *Ground*) in relation to the variety of spatial expressions in Turkish?

2. What is the relationship between the variety of locative case markers in Turkish and oculomotor characteristics, response time, and participant's utterance?

3. Is there a difference between *Core* and *Non-Core* groups, as reported in the previous work (Landau, 2017, 2020) in Turkish in terms of fixation duration and count on *Figure* and *Ground*, the response time of participants, and utterance of participants?

In Chapter 2, we present a literature review of the topic and provide the relevant background for the study. We give a brief summary of the literature on spatial language, frame of reference, Whorfian and Universalist approaches, spatial categorization, spatial term types, spatial case markers, Turkish grammar on spatial expressions, Turkish spatial language studies, and eye movement on scene viewing.

CHAPTER 2

LITERATURE REVIEW

2.1. Space and Spatial Language

Representing and reasoning about space is a crucial and fundamental ability for humans. Levinson (2003) states that the notions about spaces such as spatial scenes, spatial reasoning, spatial language including spatial terms, spatial adpositions, in particular, the idea of space itself, have been widely discussed since maybe even before the Early Greek era, from Plato to Kant. Plato saw space as air with geometrical properties, allowing tridimensionality. According to Aristotle, space was not void but a nested series of places, referring to the frame of reference in twentieth-century terms. According to Levinson, around the Renaissance, an infinite three-dimensional void started as a definition of space. Newton distinguished Absolute space and Relative space. Later on, Kant psychologizes space, claiming that absolute space organizes our perception of space. Therefore, it can be said that the notion of space itself and other notions about space are widely thought topics among scholars.

Spatial experiences and object recognition are not unique to humans; however, what makes humans special is the ability to express them. Humans have the ability to encode and speak about notions about spaces and geometrical properties around them. The linguistic representation of the spatial experiences involves three elements; the object to be located, another object for location, and their relationship (Landau & Jackendoff, 1993). There are a variety of terms used in the literature to express objects in a spatial scene. The terms *Figure* and *Ground* are used in Gestalt psychology, and Talmy (1973; 2000), Levinson (2003), and Landau (2020) use the terms to describe primary and secondary objects. Levinson also suggests the terms' theme' or 'trajector' for *Figure*, and 'relatum', 'landmark', or 'reference point' for *Ground*. Besides, Talmy also indicates that the Reference Object is symbolic usage in linguistic context to express *Ground*. In this study, mainly the terms *Figure* and *Ground* are used.

The relationship between *Ground* and *Figure* is linguistically represented mainly with adpositions. There is vast variability in spatial adpositions between languages. Moreover, there are few studies on the conceptualization of space in Turkish, so not all Turkish adpositions are covered in the literature. To decide if a lexical item in a description is a spatial adposition or not, we needed a definition. Levinson et al. (2003)

define adposition as "a spatial adposition is any expression that heads an adverbial phrase of location in the BASIC LOCATIVE CONSTRUCTION (answers to where-questions)." (p. 486). However, as Lakusta, Brucato, and Landau (2020) revealed in a recent study, the Basic Locative Construction hypothesis may not work for dynamic configurations, particularly for *Support-From-Below*. Therefore, it will be only used for terms used in static spatial scenes.

Everyday activities depend on spatial reasoning, people's understanding of spatial scenes, to be able to determine which way to go or to describe where to leave the keys. Spatial settings that humans encounter are categorized into two in the literature; dynamic spatial scenes and static spatial scenes (Talmy, 2000; Levinson, 2003). In this study, static spatial scenes are discussed only.

In summary, space is a crucial notion that has been studied over centuries. The *Figure*, the object to be located, and the *Ground,* the object for location, are the terms that are used in previous research, and in this study, we use these terms too. We also used the definition of adposition from Levinson et al. (2003) to decide whether a lexical is an adposition. In addition, spatial scenes are divided into two in the literature; dynamic and static scenes. Our study falls into static scene research. The following section presents the types of frames of reference, which are describing the same spatial disposition with different strategies, and their usage in different languages and cultures.

2.2. Frames of Reference

According to Talmy (2000), different schemas can be used for the same configuration in two ways. First, the *Ground* can have different spatial relations with several Figure objects at the same time. A book may be *on the table* while a cat may be *under the table* at the same time. Second, different schematizations can explain a spatial relation with a particular Figure and a particular Ground; that the *Figure* can be described by using ground-based, field-based, and speaker-based schematization. Which one to use is selected by speakers. However, culture or language may force speakers to use one schematization. For example, while English treats cars and buses as *Grounds* differently by requiring prepositions *in* for car and *on* for bus, Germany treats them the same. Moreover, variability in descriptions is not limited to within the same reference point; wider differences occur between languages. Wintu, the Native American language, refers to objects in earth-based geometry in contrast to the most commonly used languages in the West. For example, instead of the sentence "My right arm itches.", "My east arm itches." is standard usage in Wintu (Talmy, 2000). The differences in schematization choice may result from the language rule system or psycho-cultural effects on the language. Talmy prefers to defend the effects of culturally different emphases in schematization. Moreover, schemas are not continuous; that they are disjunctive, so language fails to be exact.

The variety of how people describe the same spatial disposition with different strategies has been discussed in the literature, mostly as frames of reference or FoR in short. From Aristotle to Newton to Kant, the idea has been discussed severely in different times and fields, such as philosophy, brain sciences, linguistics, psychology, and psycholinguistics, which used different frames of reference (Levinson, 2003). Although all FoR use a coordinate system to navigate the *Ground* and the *Figure*, Levinson claims that the underlying coordinate system they use is what makes them different (p.24). In the psychology of language, Levinson presents the renamed versions of viewer-centered frames, object-centered frames, and environment-centered frames; intrinsic, relative, and absolute (p.32). The intrinsic frame uses the object-oriented coordinate system. *Ground's* facets or sides, which are inherent features, regulate the coordinate system, as in (1):

(1) The dog (*Figure*) is in front of the house (*Ground*).

The relative frame is a ternary spatial relation with the *Figure*, the *Ground*, and a viewpoint. The coordination system sits on the viewpoint, although there may be a second coordinate system in the *Ground*. In the relative frame, the coordinate system is mostly centered on the main axis of the body, which produces opposite terms such as *up/down, back/front*, and *left/right* (Levinson,2003). The term ego-centered was also used for the relative frame by Guen (2011), as exemplified in (2):

(2) The dog (*Figure*) is on the left of the house (*Ground*).

The absolute frame's coordinate system is fixed to the *Ground*, which also explains different substitutions for the absolute frame, such as the term geo-centered by Guen (2011). Its arguments are the *Figure* and the *Ground*, as exemplified in (3):

(3) The dog (*Figure*) is in the north of the house (*Ground*).

Also, inferences of spatial descriptions only transfer across different descriptions in the absolute frame, while the relative frame is able to do in the condition of viewpoint is constant. The intrinsic frame does not have this capability at all (Levinson, 2003).

Having terms such as *front, left, back*, and *right* does not prove that the language uses a relative frame. Many languages do not have a relative frame and use these terms in the intrinsic frame. In addition, relative frames are not fundamental in the linguistic point of view, and children, in fact, become professional at using these terms at a relatively late age, which brings suspicions to the universality of the relative frame (Levinson, 2003, p. 46). Not only in the case relative frame but other frames also are not used in every language. For example, while English speakers only use the relative frame or intrinsic frame to describe small-scale spatial descriptions, speakers of Guugu Yimithirr only use the absolute frame (Majid et al., 2004). For example, while speakers of Guugu Yimithirr say, "The fly is to the north of your ears.", English speakers say, "The fly is to the left of your ears." or "The fly is beside your ears.". Therefore, some languages use only one frame, some use two, and some use all.

In sum, Levinson (2003) named the types of FoR categorizations, which are describing the same spatial disposition in different points of view as intrinsic, relative, and absolute. The FoR type affects the variety of adpositions people use in their speech. In addition, languages and cultures differ on which frames they use. While some use two types of FoR, some use just one. In our study, participants used mainly intrinsic FoR, while some used relative FoR. The preferred FoR changed in the same scene for different people, which changed the adposition they use in their utterances. Therefore, the changes in the preferred FoR are important for the variety of adpositions. The next section presents Whorfian, and Universalist approaches for spatial language, whether there is consensus on cognitive information processing of FoR between different languages, and whether differences in FoR affect speakers' cognition.

2.3.Whorfian and Universalist Approaches

Levinson et al. (2003) conducted an experiment using a commonly used battery (Bowerman, 1996; Majid, Jordan & Dunn, 2005), the topological relations picture series or TRPS for short (Bowerman & Pederson, 1992). We also used the same battery in the current study. The battery can be used for encoding prepositions, such as *on*, *in*, *under*, *over*, *near*, *against*, *inside*, *on the top of*, and *in the middle of* . The *Figure* was colored yellow and asked, "Where is [*Figure*]?", which gives basic locative constructions, as Levison et al. defined. They aimed to investigate if the close-class adpositions produce universal spatial semantics. They tested three universalist hypotheses that they think are not legitimate. First, they claimed that there was no agreement on *in* and *on* or other categories across languages. To support this hypothesis, Levinson et al. followed Mellissa Bowerman's method, and they used Venn diagrams to map adpositional groupings onto a fixed arrangement of scenes. The only grouping that languages agreed on was *under*. There was no agreement, as Levinson et al. expected. Second, They used the idea of semantic foci in colors that although lexical variability occurs between languages in color terms, orange and blue are not confused in conceptual space. They tried to show that conceptual space under topological relations is coherent. However, although some foundation was found for the second hypothesis, it was not significant. Third, multidimensional scaling was used to look at language group scenarios. Notions of *Attachment*, *Superadjacency*, *Full Containment*, *Subadjacency*, and *Proximity* were clustered cross-linguistically.

Not only Levinson et al. (2003) used the idea of semantic foci of color, but also recently, Carstensen et al. (2019) studied the spatial topological notions *in/on* in light of semantic foci categorization. Carstensen et al. (2019) conducted a cross-linguistic non-open style study with the same battery used in the current study (Bowerman et al., 1992; Levinson et al., 2003) with the speakers of Dutch, English, French, Japanese, Korean, Mandarin, Spanish. They used three characteristics of colors' foci members to assess the semantic foci of spatial topological adpositions. They asked whether there is consensus within and between languages on focal spatial relations and whether spatial categories compose the structure. To answer the questions, they asked participants to do scene naming, category naming. Later, they named the data, i.e., they chose the most used spatial term for a stimulus for encoding. This study diverges from our study at this point. Our study looked for the variety of spatial expressions at a stimulus; however, most of the studies chose to label one spatial expression for a stimulus and looked for variety between and within language.

The results of the study showed that speakers of the languages agreed on focal spatial relations, and there was the same consensus in seven languages, which showed between language similarity. Carstensen et al. (2019) also found composite structures of spatial expressions such as the Korean spatial term *wie*, a composition of *above* and *on* in English. There are similar terms in Turkish too, such as *üzerinde*, a composition of *above* and *on* in English (Atak, 2018). Cartensen et al. (2019) claimed that universal tendencies in spatial relations had been revealed empirically.

Languages use different FoR (Frames of Reference) to describe spatial dispositions, as mentioned before. While speakers of Guugu Yimithirr only use the absolute frame, English speakers use intrinsic and relative frames to say that a fork is next to the spoon (intrinsic frame) (Majid et al., 2004). In the current study, we found that in Turkish static spatial language, speakers prefer the relative frame, as exemplified in (4), and the intrinsic frame, as in (5) for the following stimulus asking, "Where is the dog?" in Figure 1.

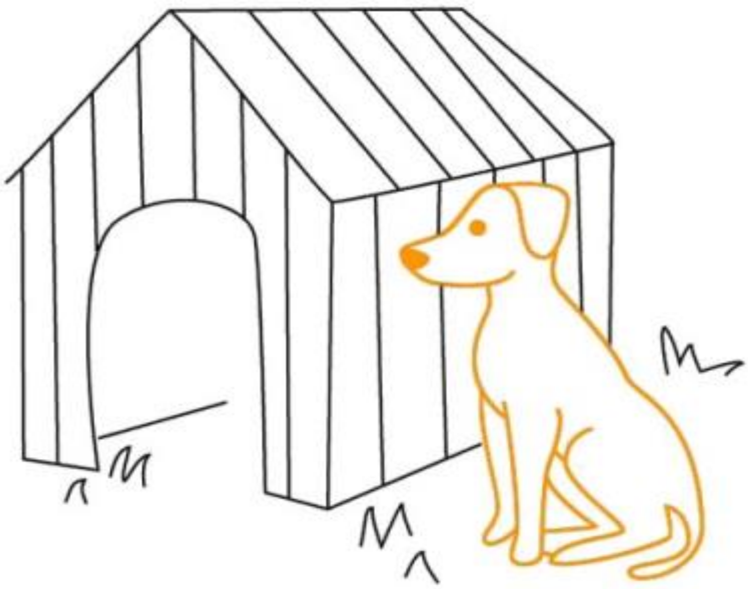

Figure 1: Stimulus depicting a dog (the Figure) near a kennel (the Ground)

(4) *kulübenin sağında*

kennel to be in the right

'on the right of the kennel'

(5) *kulübenin yanında*

kennel beside

'beside the kennel'

Turkish speakers do not use the absolute frame to describe static spatial scenes. So, does this difference in usage influence the perception of the speakers? Does the fact that Turkish speakers do not utter *in the north of the kennel* have any impact on their perception of the environment? The present study does not examine those questions; however, it has been studied before. There are two different thoughts on this. To test whether differences in frame usage affect the cognition of the speakers, various studies used non-linguistic tasks. We will show only three studies, with the one having opposite claims to the other two.

Majid et al. (2004) reported that several experiments done with non-linguistic tasks showed that speakers used the same frames they used to speak in non-linguistic tasks. In experiments, participants were shown spatial relations and then rotated 180 degrees in the opposite direction to test if they used the same FoR they usually did to solve spatial tests. One of the tasks was illustrated simply below for better understanding in Figure 2.

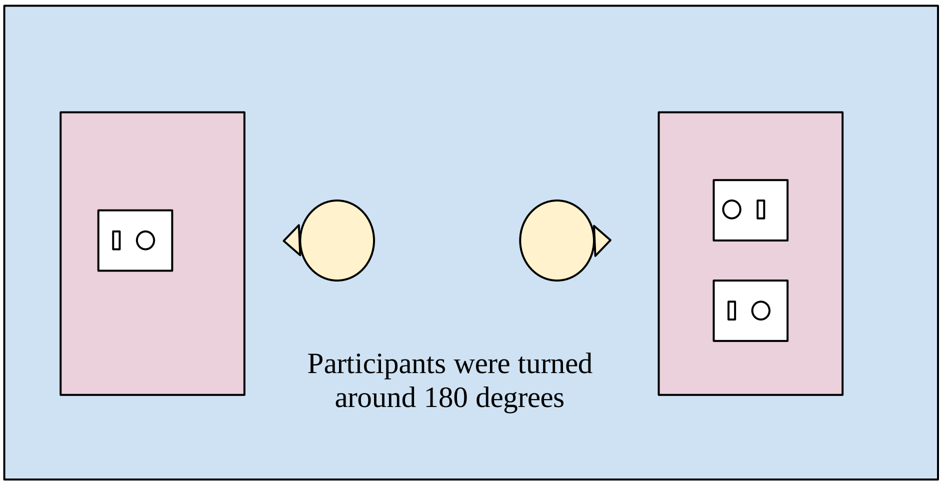

Figure 2: Setting of a non-linguistic task example

The languages that use absolute FoR, such as Arrernte, Tzeltal, Belhare (Nepal), and more, were reported to use absolute FoR in non-linguistic tasks, while languages that use relative FoR, such as Japanese, use the relative FoR in non-linguistic tasks. This parallel interaction between linguistic FoR and non-linguistic FoR was seen as an example of the Whorfian effect (Majid et al., 2014).

Contrary to studies that support the Whorfian effect in non-linguistic experiments, Yucatec Mayas have shown no preferred FoR (Le Guen, 2011). In particular, women and men show differences in absolute (geocentric as Le Guen (2011) calls) FoR; however, this does not prevent women from using absolute FoR in their gestures in non-linguistic tasks. Women and men have distinct differences in using cardinal direction because of the nature of Yutac Mayas, where men go to the forest and do the spiritual rituals that include spatial terms. In contrast, women stay inside the village and home. Four studies were conducted by Guen (2014). In the first one, people's daily lives were videotaped to record the frequencies of spatial terms used by women and men. In the second study, the linguistic knowledge of the people was assessed. Men and women were asked to talk about the objects nearby. Women had significantly less knowledge of the terms *right* and *left*. Almost 55% and 66% of women knew the terms *left* and *right*, respectively. On cardinal terms, which are used for geocentric FoR, the gender difference was huge. None of the women knew about the terms *west* and *south*, only 22% knew *east*, and 33% knew *north*. The gender difference is apparent but does it affect the non-verbal tasks, which is a question asked to understand the cognitive effect of linguistic differences among languages. The non-verbal task named *animal in a row*, a similar task explained before by Majid et al. (2004), was applied to both men and women. There were no gender differences, and both men and women chose geocentric FoR. Although women knew less about geocentric FoR, the result of the study showed that this did not affect the non-verbal task. To state which FoR Yucatec Maya use, Study 4 was, in which participants were asked to describe the location of a distant object (*Figure*) relative to others (*Ground*), was conducted for both women and men. Their gesture was analyzed along with linguistic descriptions, and it is shown that both women and men used geocentric FoR along with gestures. However, only two of the participants used cardinal terms. Results of four studies conducted by Guen showed that although women did not know linguistic cardinal terms, they were able to use geocentric FoR with their gestures, which contradicts the Whorfian claims and shows that language is not the only way to understand the structure of spatial cognition.

If speakers of a language with one preferred FoR were forced to use another FoR, what would be the result? This question was asked by Haun et al. (2011, as cited in Li & Abarbanell, 2018), who compared children speakers of Dutch, who prefer egocentric FoR, and children speakers of Hai||om, who prefers geocentric FoR. Children's ability on non-preferred FoR was assessed by the *animal on row tests*, similar to Yutac Mayas (Guen, 2011), and concluded that children have difficulty using incongruent FoR, supporting the culturally-dependent frame. However, Li and Abarbanell redesigned the experiment of Haun et al. and conducted additional experiments to minimize experimental effects on speakers with English-speaking and Tseltal speaking children. In several experiments, in line with previous findings, they found out that, when children are not instructed, i.e., in open-ended tests, they tend to apply linguistically preferred FoR. However, when children are instructed to use non-preferred FoR by minimizing experimental and participant effects, it is revealed that children who prefer linguistically egocentric FoR, such as English, and geocentric FoR, such as Tseltal Mayan, can use both systems if appropriately assessed. This shows that instruction given in the experiment plays a vital role in results. Open-ended questions without strict instruction reveal the preferred way of participants, which is the intention and the method of the current study. All in all, there is still disagreement about if FoR affects the spatial cognition of speakers and if speakers of all languages can comfortably use every FoR.

If diversity in language changes people's cognition, when does it start, and is it the only factor in spatial cognition? If not, how can differences in the usage of spatial terms be explained? Even if we are not interested in the developmental aspect of the spatial terms in children, it is inevitable while studying the topics, such as universality and *Core* spatial terms, which depend on studies on children's acquisition of spatial terms.

Empirical evidence shows that children know about space before they talk about it (Bowerman, 1996), supporting the idea that children are born with the concept of space and relationships in it. However, they cannot express it immediately, so non-linguistic cues reveal the ability of the children to understand and categorize spatial scenes, similar to women's gestures in Yucatec Mayas (Guen,2011). In addition, this ability to understand non-linguistic spatial notions affects how children learn linguistic spatial terms. Bowerman stated that E. V. Clark proposed that "prepositions whose meanings accord with learners' non-linguistic spatial strategies are acquired before prepositions whose meanings do not; hence *in* is easier than *on*, which in turn is easier than *under*."

Another claim made by Dasen (2018) recently is that using egocentric FoR or using geocentric spatial FoR depends on cognitive style. Cognitive style is "one's preferred way of processing information and dealing with tasks" (p. S94), although it might be unconscious. Even if two people have the same cognitive capacity, they may choose to use different cognitive processes which exhibit the cognitive style. Dasen studied different FoRs used by various languages in rural and urban settings, such as Bali, India, and Nepal, over the years. He concluded that all three FoRs were available for the speakers, although they preferred egocentric FoR. Another finding showed that the way cognitive tasks are applied affects geocentric FoR differently. Dasen concluded that FoR selection is a cognitive style, and language and culture override the cognition with which people are born.

In sum, there have been different opinions on spatial semantics of topological relations among languages between Whorfian and Universalist approaches. Levison et al. (2003) showed that close-class adpositions do not produce universal spatial semantics. In contrast, Cartensen et al. (2019) used the same idea and battery, which we used in the present study, and concluded that there was a consensus between languages. The different approaches are also tested on the preferred FoR of languages and their cognitive effects. Majid et al. (2004) showed that when speakers prefer absolute FoR in linguistic tasks, they prefer the same FoR in non-linguistic tasks, which implies the language-dependent effect of FoR on cognition. However, Le Guen (2011) showed that although women of Yucatec Mayas did not use geocentric FoR in their speech, they used it in the non-linguistic task. Moreover, Haun et al. (2011, as cited in Li & Abarbanell, 2018) forced children with different preferred FoR to use another FoR and concluded that they could use both if appropriately instructed. These findings of Le Guen and Haun et al. support Universalist approaches on FoR. Moreover, Dasen (2018) claimed that FoR selection is preferred, all three FoRs are available for speakers, and language and culture supersede the cognition they are born with.

The current study uses the same battery Levinson et al. (2003) uses; however, the aim of the study is not to refute Universality or the Whorfian hypothesis of spatial terms. We accepted the universalist claim of Landau (2017, 2020) on *Core* and *Non-Core* agreement across languages to label the stimuli as *Core* and *Non-Core* in the analyses. Moreover, we analyzed various spatial expressions at a stimulus; however, most of the studies chose to label one spatial expression for a stimulus and looked for variety between and within language. Moreover, in the present study, participants did not use the absolute frame to describe static spatial scenes. Details of the FoR used in Turkish are explained before in this section. The following section presents cross-linguistic perspectives on spatial categorization and how we can interpret these differences in Whorfian and Universalist approaches.

2.4. Cross-linguistic Perspectives on Spatial Categorization

If children have pre-linguistic spatial concepts, how do they group them, and does it affect the linguistic spatial concept groupings? In terms of spatial concept groupings, languages vary considerably. For example, English, Finnish, Dutch and Spanish groups contact with and support by a vertical space, which Landau (2020) considers as *Non-Core*, and contact with and support by a horizontal space, which Landau (2020) considers as *Core*, and *Containment*, which Landau (2020) also considers as *Core*, differently. In Spanish, all of them are put in one group and described by spatial terms *en*, while Dutch separate them all and describe them with spatial terms *aan*, *op*, and *'n*, respectively. In addition to western languages, the range of variety gets more comprehensive in the other languages, such as Tzeltal, which prefers absolute FoR, which is not common among the languages mentioned before (Majid et al., 2014). Tzeltal has seven different words to replace *in*, such as *t'umul*, which means to be located in liquid (Bowerman, 1996). Therefore, it seems that different languages have different linguistic spatial concept groupings. Moreover, despite the richness of babies' non-linguistic spatial concepts, the Whorfian camp is not entirely impossible considering that children have to learn languages that vary considerably.

Recently, Yun and Choi (2018) conducted an experiment to support the rising idea that neither the Universalist camp nor the Whorfian camp is mistaken; the relationship between language and cognition is just complex. Native Korean and English speakers described dynamic spatial scenes, and non-linguistic data categorizations from Choi and Hattrup (2012) were assessed against linguistic data. Clustering and MDS analysis have shown that while speakers of Korean used diverse verbs, speakers of English did not use verbs; instead, they used limited close class prepositions. In addition, it was known that the sensitivity of speakers to categories was different for English and Korean; they both showed additional sensitivity to other categories. *Containment* and *Support* sensitivity, represented by *in/on* in English, could be found in Korean speakers, although speakers of Korean show sensitivity to the *Tight* and the *Loose* category. The opposite applies to English and, both languages showed sensitivity to the verticality of the objects. Moreover, the categories of *Loose-Support* and *Tight-Support* were expressed with similar adpositions in the languages, but expressions differ in non-linguistic tasks. In sum, the analysis and results showed that language data predict non-linguistic data. Therefore, spatial cognition is influenced by spatial semantics, but both have independent areas. In short, Yun and Choi claimed that commonalities and differences of Korean and English speakers showed that universal perception was the base of spatial cognition; however, each language structures its own semantic system on a universal basis, distinguishing languages from each other.

Recent comparative studies on spatial categories support the idea of language structure addition to universal core semantics (Yun & Choi, 2018). Feist and Zhang (2019) conducted a comparative study of English and Mandarin. They used multidimensional scaling to produce similarity space, a visual representation of spatial terms closeness in conceptual space. To illustrate, if both spatial relationships are expressed with *on*, they are placed close to each other in the similarity space. Instead of comparing each scene, the conceptual spaces were compared. The analysis has shown that languages vary in spatial semantics complexity, and Feist and Zhang claimed that complexity is combined to core universal concepts.

In sum, spatial concept groupings vary depending on the languages, which seems to result from language effect over non-linguistic spatial concepts of babies. This relationship between language and cognition is recently considered that language-specific semantic system is structured on the universal perception of spatial cognition. In the next section, we present satellite-framed and verb-framed languages.

2.5. Satellite-framed and Verb-Framed

Motion is represented typologically in two different ways in languages; satellite-framed and verb-framed (Furman et al., 2010; Talmy, 2000; Toplu,2011). Like most Indo-European languages, English is a satellite-framed language whose verbs have satellites that express the path. In addition, Furman et al. also use the path as the relation of the motion. For example, in the verb phrase 'go up', 'go' tells what the subject does, and 'up' gives the subject direction. In contrast, in verb-framed languages, the path is presented inside the semantically specific verb. For example, 'çıkmak' in Turkish has the same meaning as 'go up'; however, the path is inside the verb. Turkish is a verb-framed language (Furman, Özyürek & Küntay, 2010), similar to Hebrew, Spanish, and Korean. In addition, semantically general verbs are also quite common, along with semantically specific words. For example, *koymak* 'put', which is semantically general, and *sokmak* 'put in', which is semantically specific, can both be used in the sentence 'The girl put the book in her bag.'.

Even though the current thesis studies static spatial language, there were descriptions in some stimuli that do not fit the static spatial language frame in Turkish. While English speakers would need satellites, Turkish speakers used verbs semantically specific verbs in the study, such as *askıya asılı* 'hanger-DAT hanging on'.

The next section presents spatial term types and *Core* and *Non-Core* distinction by Landau. (2017, 2020).

2.6. Spatial Term Types and Core vs. Non-Core Subtypes

Objects are categorized by their distinct geometric properties. However, when they are represented as the *Figure* and the *Ground*, most of the geometric distinctions of the objects are disregarded. Objects lose their most geometric structure when given roles of the *Figure* and the *Ground* in English. Although there are other languages regarding geometric properties of more than English, they still do not encode as much as object names (Landau & Jackendoff, 1993). According to Landau and Jackendoff, there is a distinct division between *what* system and *where* system. An object is described by geometric properties, i.e., complex representations of shapes and surface, however expressing where the object only needs key properties, axial structures, leaving the geometric details, which is described by close-class spatial expressions. The *Figure* needs no geometric requirements, and the *Ground* needs a little more than the *Figure*, yet, very limited. The distinction of the *what* and *where* systems are also encoded in the neural system. The ventral stream carries out tasks involving *what* system, and the dorsal stream carries out tasks involving *where* system. When animals with a lesion to ventral stream are given tasks about object shape, they could not carry out; however, they could respond to object location. Similarly, animals with a lesion to the dorsal stream respond the opposite way (Landau, 2017).

Landau (2017) later proposed a second division for *where* system. Spatial prepositions were updated and divided into functional and geometrical groups. The first group includes prepositions *on/in*, named *topological terms*, representing force-dynamic and mechanical meanings. They are claimed to represent the properties of openness/closeness and interior/exterior rather than the properties of Euclidian geometry, such as angles and distance. The second group represents a geometrical relationship between groups and includes prepositions *below/above*, *right/left*, and *east/west*. For the first group, other factors rather than restricted geometrical properties of the objects might be involved in the semantics of the preposition. For example, *on* express support from the surface, however, "fly on the ceiling" holds a force-dynamic relationship between the fly and the ceiling. There is no support from the ceiling to the fly. The relationship is functional, not geometrical. Therefore, the nature of the *Figure* and the *Ground* changes the use of *in/on*. Another example from Vandeolise (1991, as cited in Landau, 2017) is how words *bowl* and *tray* are used because of functional differences. Although curved tray and shallow bowl have similar geometrical properties, one says "apple in the bowl" and "apple on the tray". How the *Ground* is perceived, rather than geometrical properties, affects the preposition selection. The way the *Ground* is categorized changes between cultures and languages. Despite similarities between cultures on the contents of *Containment* or *Support*, there are pretty clear differences between cultures on categorizing the *Ground*. Hence, the function-based prepositions cause cross-linguistic variation. In addition, the fact that mechanical relationships between objects are infinitely complex makes learning particular relationships difficult.

On the other hand, geometry-based prepositions, engaging distance, and direction present an entirely different pattern. To illustrate, *up, down/above, below/right, left/north, south, east, west, uphill/downhill* express direction and distance, unlike *in/on*, and because direction and distance are concepts represented before language as infants as early as 3–4-month-old, geometry-based prepositions display consistency, and little-variability across-languages (Landau, 2017). Not only in English but other languages also support the claims of Landau. Recently, Žilinskaitė-Šinkūnienė et al. (2019) analyzed the Baltic languages and Estonian in a production task to review if claims of Landau can be proven across languages. They concluded that the *Containment* and the *Support* categories are comprehensive and complex, and the consistency holds most for *Core* subtypes. In addition, they supported Landau that the differences between languages come from the functionality nature of *Containment* and *Support*. They also reported that in their previous research with the geometric framework (RCC+F) on *Containment* and *Support* categories, the Baltic languages constrained geometrically and showed limited variation (Žilinskaitė-Šinkūnienė et al. 2019). These findings support the claims of Landau on differences between function-based and geometry-based spatial prepositions.

Johannes, Wang, and Landau (2015) discussed the Whorfian effect in cognition, using the term cultural-dependent frame, which is the opposite of the universalist frame. In the culture-dependent frame, it is claimed that culture should be understood completely to understand spatial languages and terms, even *Core* ones. In contrast, the universalist frame claims that *Core* terms should be encoded similarly across languages. Levinson et al. (2003) and Rigier et al. (2013), siding cultural-dependent frame, both used the same battery the current study uses (Bowerman & Pederson, 1993). They showed that the usage of adpositions varies across languages, and spatial encoding systems are complex and different from each other. In contrast to the cultural dependent frame, the universalists frame camp tried disparate measures that have not been used in cross-language spatial term research.

To show that the cross-linguistic differences of function-based prepositions are not necessarily a drawback from universal spatial cognition, Landau et al. (2017) studied with English and Greek 4-year-old children and adults to show that both English and Greek have *Containment* and *Support* types that are divided into subtypes, including

Core ones. *Containment* and *Support* are types of spatial configurations. *Containment* includes subtypes of *Full Containment- Loose Fit, Full Containment- Tight Fit, Partial Containment- Loose Fit, Partial Containment- Tight Fit, Interlocking, and Embedded*. *Support* includes *Gravitational Support, Embedded, Adhesion, Hanging, Point-attachment* subtypes. *Full & Loose, Full & Tight, Partial & Loose*, and *Partial & Tight* subtypes were assumed to be *Core* in the *Containment* category. In the *Support* category, the *Gravitational Support* subtype was considered to be *Core*. The rest was assumed to be *Non-Core* (Landau et al., 2017; Landau, 2020). For example, a representation of 'crack in a mug' is included in a subtype called *Embedded* of *Containment* type. The *Embedded* subtype is considered in *the Non-Core* group. However, a representation of 'apple in a bowl' is included in a subtype called *Full Containment-Loose Fit* of *Containment* type. *Full Containment-Loose Fit* is considered to be in the *Core* group.

The purpose of the research of Landau et al. (2017) was to study the existing belief that humans put data on the existing pre-linguistic spatial notion to learn spatial expressions. They created a real-life object battery, consisting of *Containment* and *Support* type spatial configurations, and asked the question, "Where is X?", where X is the *Figure,* to obtain answers in basic structure construction. They analyzed the differences and commonalities within and between languages.

The most important finding for the current study is that for the *Containment* category, both adults and children used *BE in* more for *Core* subtypes. Among the subtypes, while adults preferred *Full Containment* over Partial Tight to use *BE in* more, unlike adults, children preferred Loose over Tight. Therefore, Landau et al. (2017) claimed that the sensitivity of the *Loose/Tight* category of children might indicate that this preference might be seen in other languages, which was seen in Korean. Instead of the *Containment* and *Support* category, speakers of Korean shows sensitivity to *Loose* and *Tight-fit* categories (Yun & Choi, 2018).

For the *Support* category, both children and adults used *BE on* in their *Core* subtypes except children speakers of English. The reason for English-speaking children to use *BE on* in every subtype is explained by Landau et al. (2017) as children's lack of linguistic resources. Johanson and Papafragou (2014) made a similar explanation for overextension of children of *Containment* type to *Cover* and *Occlusion* types considering conceptual overlap hypothesis. In addition, in some scenes, adult speakers used lexical verbs instead of the basic expression *BE in*. For example, instead of "the hole is in the sock", speakers used "the sock tore" in English. These usages were not common in Greek. Landau et al. explain this with English being satellite frame and Greek being verb framed; however, Turkish speakers also use such expressions despite its verb framed nature. Even though we did not include it in the analysis, participants produced such utterances. Overall, Landau et al. concluded that despite differences between Greek and English encodings, they both showed similar divisions between subtypes, which shows the parallel structure between languages.

In a recent study, Lucasto, Brucato, and Landau (2020) also tested a part of the same hypothesis, i.e., the *Core* type of *Support* is represented by the *Support-From-Below* relationship with younger children aged less than 4 years old. In the study, young childrens' spontaneous language productions were coded, so although it was closer to real-life situations, the descriptions mainly were dynamic spatial scenes. Use of partner verb and *on* of children, aged less than 4 years old, have shown that they also distinguish *Core* group of the *Support* type, which is *Support-From-Below*, such as 'cup on the table.' *BE on* was used mostly for *Support-From-Below* configurations, i.e., *Core*, whereas lexically rich verbs, such as *hang* and *stick*, were used for mechanical support, i.e., *Non-Core*. For example, the relationship between a coat (*Figure*) and a hook (*Ground*) is considered *Non-Core* by Landau et al. (2017), and children expressed the relationship as "Coat is hanged on the hook.". They used the lexically rich verb "to hang on" instead of "*BE on*".

Although young children's preference implies that they are predisposed to *Core* notions, it is also revealed by parents' data that children are sensitive to how parents use verbs in *Non-Core* spatial configurations. Therefore, language, children exposed, may also be an effect.

Similar to Landau et al. (2017), Johannes et al. (2015) hypothesized that the usage of basic expression *BE in/on* in subtypes of *Containment* and *Support* is similar in different languages. Speakers of typologically different three languages, i.e., English (Germanic), Hindi (Indo-Iranian), and Mandarin (Sino-Tibetian), were asked to describe 80 static spatial scenes to analyze if there are within language variations that indicate similarities across languages on the way they encode spatial relationship. Johannes et al. suggested that assessing within-language variation in spatial expressions can give cross-linguistic similarities. Hence, they studied two types of within language variation. First, the systematicity in basic expression use in each language was analyzed and found that commonalities were found in basic expression usage among subtypes of *Containment* and *Support* categories. Second, Johannes et al. hypothesized that speakers of different languages should display similar variations in the number of expressions for different types of relations. Similarity, entropy, which provided data about variability and frequency of expressions for each scene, was calculated across scenes in each language. Few studies in the literature carried out analyses on variation in the spatial expression encodings for each spatial relation scene. In studies on spatial relations, in general, the most used spatial expression is labeled to a spatial scene, and variation between categories of spatial relation types, such as *Containment* and *Occlusion*, or variation between languages are assessed.

Nevertheless, similar to our study, in the second part of the study, Johannes et al. (2015) analyzed how speakers of languages preferred to use either frequently used few expressions or many low-frequency expressions. The variation of spatial expressions of spatial categories and the correlation between languages have been found. Overall, despite typological and cultural differences of languages, extensions of basic spatial expressions and variations in spatial scenes showed similarities.

Barbara Landau (2020) elaborates the notion of *Core* and claims that children have non-linguistic knowledge about *Core* spatial concepts of *Containment* and *Support* subtype because *Core* and *Non-Core* spatial concepts are conceptually different. While the plain verb *BE in/on* or corresponding in other languages are used in *Core* spatial concepts, information-rich lexical verbs and spatial terms are used for *Non-Core* concepts. According to Landau, the reason why differences between *Core* and *Non-Core* notions are explicit is that *Containment* and *Support* relations give the information of an object relative to others and gives the mechanical relationship of the objects so *Containment* and *Support* relationships use Topological geometry rather than Euclidian geometry. In particular, Landau and Jackendoff (1993) propose that spatial propositions are divided into two: geometric and functional. Geometric spatial prepositions reveal distance and direction, properties of Euclidian geometry, within a reference system such as *above, below, left, and right*, which has limited variation cross-linguistically. However, functional spatial propositions, such as *in* and *on*, give the mechanical information of objects and relations, properties of topological geometry, which provide variability across languages (Lakusta, Brucato & Landau, 2020). Therefore, Landau states that the claims defending the Whorfian effect by stating there is no cross-linguistic agreement on *in* and *on* (Levinson, 2000a; Levinson, 2003) should consider that the cross-linguistic variation comes from the mechanical nature of functional spatial terms. Although studies vary on the frequency order of *BE in/on* usage across *Core* subtypes of *Containment* and *Support* types, it is shown that *BE in/on* was used the most in *Core* subtypes on which Landau et al. (2017) decided.

Overall, L&J (1993) distinguished what system; requiring geometric properties and where system, represented by the *Figure* and the *Ground* with limited geometrical properties. Later on, Landau (2017) divided where system too; prepositions with force-dynamical relationship and prepositions with a geometrical relationship. The former represents complex and indefinite mechanical relationships. The latter engages universal distance and direction, so while the latter is consistent among languages, the former is not. This claim is supported by different language families other than Germanic (Žilinskaitė-Šinkūnienė et al. 2019). Therefore, Landau (2017) claimed the differences on *in/on* between languages come from force-dynamic nature. In addition, Landau et al. (2017) divided force-dynamic prepositions into *Core* and *Non-Core* groups. They claimed that although languages vary on force-dynamic spatial terms, they show similar usage of basic expressions in subtypes. In addition, Johannes et al. (2015) claimed that cross-linguistic similarities could be found in within-language variations.

In the current study, In the light of studies of Landau et al. (2017) and studies on *Core* vs. *Non-Core* subtypes (Johannes et al., 2015; Landau, 2020), we specified *Containment* and *Support* types and, later, divided these types into *Core* and *Non-Core* groups in the stimuli, TRPS. We used the full stimulus set presented in the appendix section of Landau et al. for categorization. Instead of a cross-language study, the examination of behavioral differences of participants between *Core* and *Non-Core* subtypes was conducted to test the universal *Core* and *Non-Core* distinction of Landau et al. in Turkish by assessing within-language variation of spatial expressions. A similar method was used by Johannes et al., in which the variation of spatial expressions of spatial categories was analyzed. The following section presents Turkish grammar and the special case markers related to spatial expressions.

2.7. Locatives and Turkish Grammar

The relationship between the *Figure* and the *Ground* is not described only by adpositions. Along with adpositions and verbs, case markers are used by some languages to express spatial relationships between objects. Turkish is one of the languages that use locative case makers for spatial description. However, spatial language is analyzed mainly through adpositions in studies, even if the studied language uses case markers to express spatial relationships (Atak,2018). Therefore, there are a limited number of studies on the relationship between locative case markers and spatial cognition in the literature.

The Baltic languages are one of the groups that use locative case markers to express spatial relationships. Recently, Žilinskaitė-Šinkūnienė et al. (2019) described how subtypes of *Containment*, i.e., full and partial *Containment*, *Interlocking*, and subtypes of *Support*, i.e., *Support-From-Below, Adhesion, Hanging* and *Encirclement-with-Contact*, are expressed in the Baltic languages, namely, Lithuanian, and Latvian, and in Estonian, which are spoken in the same area. While Estonian is considered between verb-framed and satellite-framed languages according to Talmian lexicalization patterns, Lithuanian and Latvian are considered satellite-framed languages. Lithuanian and Latvian use one locative case to express both small and large-scale environments. Baltic languages also use adverbial elements, in addition to adpositions, to describe spatial scenes. To analyze semantic categories of *Containment* and *Support* relations in Lithuanian, Latvian, and Estonian, Žilinskaitė-Šinkūnienė et al. used an open-ended production task, similar to our study. Using open-ended tasks allows participants to use any spatial expressions they want by preventing lexical influence, resulting in case markers in the data. However, the frequency of locative cases, not the variety, in each subtype were computed, which diverge from our study.

The result of the analysis showed that in the *Containment* subtype, speakers of Lithuanian and Latvian used the locative case most of the time. Distribution over subtypes decreases from *Full Containment* to *Partial Containment* to *Interlocking* in both languages. In *Support* subtypes, locative case markers are used less in both languages. In the *Core* subtype of Lithuanian and Latvian, adpositions corresponding *on* are used almost all the time. The locative case is used in 1% of expressions in Lithuanian and in 5% of expressions in Latvian. For *Non-Core* subtypes, *Adhesion, Hanging,* and *Encirclement with Contact*, expressions exhibit variations within and between languages, with small percentages of locative cases except the *Hanging* subtype of Latvian. In sum, locative case markers are used in *Containment* and *Support* types; however, while the frequency in *Support* type is low, the frequency in *Containment* type is relatively high, with the *Core* subtype having the highest. (Žilinskaitė-Šinkūnienė et al., 2019)

In Turkish, the speakers use two ways to describe the relationship between the *Figure* and the *Ground*: adpositions and general postpositional locative case markers (-de/-da) suffixed to the *Ground* (Sümer et al., 2012). Postpositions are the nouns by addition of possessive and case markers (Göksel & Kerslake, 2005), as in (6):

(6) *ön-ün-de*

'front-POSS-LOC'

The current study takes both types as spatial terms. The locative case markers do not reveal the exact nature of relationships given by adpositions made up of lexemes and locative case markers. For example, The English sentence 'The dog is in the house.' can be expressed in Turkish in two ways; Turkish expression with locative case marker, as exemplified in (7), and Turkish expression with adposition as in (8):

(7)*Ev-de köpek var*.

'House-LOC dog there is'

(8) *Ev-in iç-i-nde köpek var.*

'House-POSS interior-GEN-LOC dog there is'

Examples from the current study for practicing two types in one spatial configuration can be seen for stimulus 11, see Figure 3.

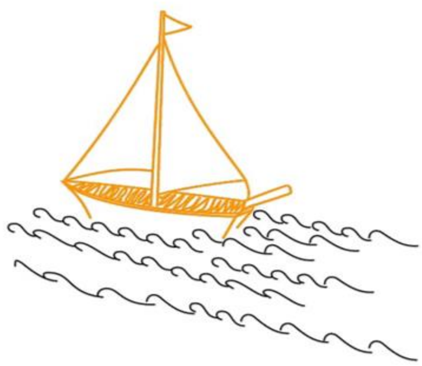

Figure 3: Stimulus 11, depicting ship (the Figure) on the sea (the Ground)

Description with Turkish locative case marker, *deniz-de* ‘sea-LOC’ and description with Turkish adposition, *deniz-in üst-ün-de* ‘sea-POSS top+GEN+LOC’ was produced for stimulus 11.

There are three types of case suffixes in Turkish that are used in spatial configurations; dative case -(y)A (*dere-ye* ‘river-(in)to), locative case -DA (*ev-de* ‘home-at’), and ablative case -DAn (*okul-dan* ‘school-from/of/out of’) (Göksel & Kerslake, 2005). In the static spatial scenes, locative case markers are used.

However, there were descriptions in the current study that use dative cases, for example, duvar-*a* ‘wall+DAT’ ‘to the wall’

Postpositions in static spatial relations are compositions of a noun, possessive suffixes, and locative case suffixes. For example, *ön-ün-de* ‘front + 3SG.POSS + LOC’

A set of Turkish spatial relation nouns are in Table 1 with (approximately) corresponding English terms (Göksel & Kerslake, 2005).

Table 1

Some Turkish Spatial Relation Nouns

Turkish Noun	English Noun	English Adposition
Ön	Front	In front of
Arka	Back	Behind
İç	Interior	Inside/in
Dış	Exterior	Outside
Üst	Top	On the top of/above/on
Alt	Bottom	Under/underneath/below
Yan	Side	Beside/next to
Karşı	Opposite side	Opposite
Ara	Space	Between/among
Etraf/çevre	Surroundings	Around
Öte	Farside	Beyond

This section presents examples from languages that use case markers and Turkish grammar related to spatial language. In Turkish, spatial configurations are expressed with nouns by using the case marker and adpositions. The following section presents a summary of studies conducted on Turkish spatial language and cognition.

2.8.Studies on Turkish Spatial Languages

The literature on Turkish spatial language is limited and not comprehensive. Turkish spatial language studies mainly consist of children's language development, spatial term usage frequencies, and motion event studies that emphasize the verb-framed nature of Turkish.

Atak (2018) studied the frequency of the spatial terms used in a task and preferred frame of reference. He combined product method and acceptability rating tasks by Carlson and Hill (2007) in four tasks. Participants were asked to fill the empty place with spatial terms in the sentence describing the *Figure* according to the *Ground*. The *Ground* was placed in the middle of a 9x9 or 7x7 frame. The *Figure* was placed in various cells in each stimulus. He mapped the most used spatial terms by participants in every cell. He also showed that Turkish speakers do not use absolute FoR, and they use intrinsic and relative FoR. In addition, the term *üzerinde* 'on the top of/above/on' is both used for *Support with Contact* and the area below the *Ground*. In addition, Atak and Günay (2012, as cited in Atak, 2018) conducted a survey of 65 *Ground* and *Figure* configurations in Turkish to test the acceptability rates of the sentences to show that the geometrical features of objects affect which object to be the *Ground* or the *Figure*. To illustrate, while the acceptance ratings of the sentence, *Bisiklet kamyonun önünde.* 'The bicycle is in front of the truck', was 94%, the acceptability rate of the sentence *Kamyon bisikletin arkasında.* 'The truck is behind the bicycle.', was 43%. On the whole, Atak provided spatial terms maps and FoR preferences of Turkish speakers in literature according to different *Figure* and *Ground* situations; however, the methodology of the study measures only the frequencies and does not involve measures such as response time or behavior variables.

There have been cross-linguistic studies, including Turkish, mainly in the area of children's language development. Johnson and Slobin (1979, cited in Atak, 2018) analyzed the development of the spatial terms of children between age 2 and 4 *in* English, Italian, Croatian, and Turkish. They concluded that the development of *in, on, under,* and *besides* is completed first. *Between, back, front* comes the next, and *back, front* comes the last in the development order. Turkish is a postpositional language, and it has less variety in spatial terms. According to Johnson and Slobin, postpositional languages are learned easier, and variety in spatial terms hinders learning. However, the example given for the differences between varieties lacks knowledge in Turkish. The problem of cross-linguistic studies, in general, is that the spatial terms in Turkish are not included in the studies efficiently enough.

Similarly, Johanson and Papafragou (2014) conducted two frequency studies, investigating why children overextend *Containment* expressions, such as *in* in English and *mesa* in Greek. Children use *Containment* expressions for expressions of *Cover* and *Occlusion* on the grounds that the way children model pre-linguistic spatial scenes can be inferred by investigating the overextension of children. In the first study, Johanson and Papafragou analyzed the overextension in *Containment* expressions of children learning Greek, a satellite-framed language, and English, a verb-framed

language, between age 4 and 5 in motion events. They concluded that frequency, user, and motivation predictions for the conceptual overlap hypothesis fit the extension. In addition, infants can distinguish spatial scenes of *Containment*, *Occlusion*, and *Cover* from an early age. Hence, the reason why spatial expressions for *Containment* scenes are used for *Occlusion* and *Cover* is not a deficiency of conceptual understanding, instead, overlap between scenes and lack of word knowledge. In the second study, ten different languages and Turkish were included. It is interpreted that the conceptual overlap hypothesis also holds cross-languages, and the bias of children shows that *in/on/under* might be under the same *Containment* type. Consequently, universal biases of mind and cognition can be interpreted by universal consistency of metaphorical extension across languages.

In the second part of the study, Turkish adpositions were analyzed. In the Turkish *Containment* data, given in tables, adpositions and case suffixes were not used efficiently. For non-verb expressions, only *içinden* 'from its inside' and *içine* 'to its inside' were used. However, Turkish speakers use other adpositions such as *dışına* 'to its outside' in this kind of spatial scene, which was not mentioned in the study. In addition, the study did not involve dative and ablative case suffixes, which are used in the dynamic spatial scenes in Turkish. The reason for this might be the lack of the number of participants, which is two in Turkish, for each language. The number of speakers in each language was either one or two, which may impact the result.

Toplu (2011) also had a cross-linguistic study with speakers of Turkish and French, which are verb-framed languages, and English, which is satellite-framed language to challenge the linguistic relativity hypothesis, i.e., Whorfian effect. Motion events were used in the study. Two non-linguistic, one linguistic, and one eye-tracking task were applied. Verbal data was in line with Talmyan typology. In addition, non-linguistic data gave uniform patterns in three languages in contrast to the linguistic relativity hypothesis.

Arık (2017) asked how speakers of Turkish encode place, motion, and orientation in motion macro-verbs. In the paper, verbs in spatial language were grouped as follows[2]:

---

[2] The translations are approximate and translated by the author.

1) Stative Verbs

   a) Angular Static: (*solunda* 'left' / *sağında* 'right' / *önünde* 'front' / *arkasında* 'behind' *dur*- 'stand' & *otur*- 'sit')

   b) Topological Static: (*içinde* 'inside' / *üstünde* 'below & on' / *altında* 'under' *dur*- 'stand')

2) Dynamic Verbs

   a) Common Dynamical (*git*- 'go' & *koş*- 'run')

      i. Only Motion Path (*çık*- 'go up' & *yaklaş*- 'come close')

      ii. Only Motion Manner (*yürü*- 'walk' & *tırman*- 'climb')

      iii. Motion path and manner (*yuvarlanarak düş*- 'fall over')

   b) Causal Dynamical (*itmek* 'push' & *ittirmek* 'cause someone push' & *çarpmak* 'hit')

      i. Motion starter

      ii. Motion stopper

It was concluded that verb-framed languages do not have to encode manners of motion. In addition, in the scenes including both motion path and manner, motion manner was encoded more. In addition, in the scenes including path, verbs with only motion path were always encoded. It was also concluded that the differences between languages were not cognitive; instead, it was because of the context (Arık, 2017).

In addition to dynamic spatial scene studies, Furman, Özyürek, and Küntay (2010) studied how children describe caused motion events and if language-specific tendencies in placement events are subtypes of caused-motion events, extend to whole caused motion events. It is found that in motion event descriptions, only verbs were used by Turkish-speaking children at first, which is also the case in other verb-framed languages. Later on, they talk about semantic elements but continued using the verb-only structure, and age did not affect the usage of semantically specific and general verbs. The fact that adults also used verb-only expressions showed that verb-only expressions were not used because of developmental reasons.

Arık (2003, 2009, as cited in Atak, 2018) also studied the spatial language of Turkish Sign Language. (TSL in short) and spatial language in spoken language. In addition to findings on TSL, it was found that while participants of spoken language preferred intrinsic FoR, participants of TSL preferred both intrinsic and relative FoR.

Similar to Arık, Sümer et al. (2012) compared spatial language in TSL and spoken language, particularly locative expressions, using picture description tasks. They used a battery containing spatial configurations in which the spatial terms *in, on*, and *under* were used. It was revealed that hearing adult participants preferred spatial lexemes more than locative case markers. Locative case markers were preferred rarely, and there were no significant differences between adults and children. To our knowledge, this is one of the few studies in the literature that analyzes locative suffixes in the Turkish spatial language, along with the current study. However, the result of Sümer et al. and the present study does not match with each other. Locative suffixes appear hugely in the utterance of participants in the present study, as seen in the result section.

Further, Sümer et al. (2014) compared TSL and Turkish spoken language for view-point dependent relations, learned by children later compared to other FoR. It is found that adult speakers of spoken language use either general relational terms such as *yanında* 'at the side of' or viewpoint-dependent lexeme such as *solunda* 'left of' for view-point dependent relations. Moreover, it was revealed that spatial terms *front* and *behind* are produced earlier than spatial terms *left* and *right* by Turkish-speaking children.

In sum, Turkish spatial language and cognition have been studied with children's language development, spatial term usage frequencies, and motion event studies. Atak (2018) studied the adposition frequencies according to the *Figure*'s spatial orientation and showed that Turkish speakers do not use absolute FoR. Moreover, it is found that children learn *in, on, under*, and *besides* first, *between, back, front* second, and *back, front* lastly (Slobin, 1979, as cited in Atak, 2018). Research with children showed that universal biases of mind and cognition could be interpreted by universal consistency of metaphorical extension across languages (Johanson & Papafragou, 2014). One cross-linguistic study, including Turkish, found consistency between English, Turkish, and French in expressions of motion events (Toplu, 2011). Several studies on motion event descriptions and findings showed that verb-framed languages do not have to encode manners of motion, and age does not affect the usage of semantically specific and general verbs. In addition, studies with TSL showed that hearing adult participants preferred spatial lexemes more than locative case markers.

The current study is similar to Atak (2018) in a way that both analyzes Turkish static spatial terms. While Atak examined frequencies of used spatial terms, FoR, and gender differences, the current study examined various spatial terms and locative suffixes and their relationship with behavioral data. Atak did not provide any behavioral data of the participants and conducted the study with restricting instructions, only letting lexemes for spatial terms and providing words for the *Ground* and the *Figure*, whereas the current study used open-ended method, letting participants choose lexemes for the *Ground* and the *Figure* and allowed them to use locative case markers. In addition to locative case markers, the current thesis analyzed the data by examining *Core* and *Non-Core* groups of spatial terms.

The next chapter presents eye movement methodology, how fixation measurements can be interpreted, and its cognitive implications in scene viewing.

2.9. Eye Movements and Scene Viewing

Recoding the eye movements of humans has been used to explain various topics from psychology to cognitive science to computer science. What people find interesting, where people direct their attention, and how they perceive what they see, have been

studied through the eye movements of people for almost a century (Duchowski, 2017). Many facts about eye movements have been discovered with research on reading with eye movements between the 19th and 20th centuries. Afterward, the behaviorist movement in experimental psychology drew attention from cognitive processes

research to applied research. This approach was held until the mid-1970s, when new technology emerged. Advances in eye-tracking let researchers compute with extensive data and produce new techniques (Rayner, 1998). Since then, research on the various topic have been conducted with eye-tracking methodology. However, most eye-tracking research investigated people's reading patterns and behaviors, which is easier to follow. However, scene viewing has no apparent patterns for particular objects, so studies have varied on what patterns, numbers, and duration of eye movements tell about cognition (Duchowski, 2017).

How does the semantic of the static scene affect the placement and duration of the eye fixation on the scenes? Answers to this question contribute to research on the online measurement of visual and cognitive information processing (Henderson & Hollingworth, 1998; Henderson, 2003). Therefore, eye fixation is an informative variable for visual cognition in scene perception. Observers' eyes move rapidly (saccades) between relatively stable moments (fixations) while looking at a static scene (Barthelme et al., 2013). The fixated eyes are not really still because of tremors of the eyes (Rayner, 1998). Moreover, fixations, where our eyes are pointed, give us information about what people see, remember, and understand on a scene. Ongoing perceptual, cognitive, and behavioral activities cause people to fixate on particular places (Duchowski,2017). Visual attention is also related to eye movements. (Henderson, 2011). Therefore, eye-movement research is essential to understand scene perception fully (Rayner & Pollatsek, as cited in Duchowski, 2017), and visual and cognitive information processing can be measured by using eye movement data (Henderson & Hollingworth, as cited in Duchowski, 2017).

One question on eye movement control is the degree to which it is controlled directly. Previous studies have shown that a proportion of fixations remain indirectly controlled. DeGraef et al. (1990, as cited in Rayner, 1998) conducted an object decision task in which participants detect non-object in the scene. They found that scene context affected fixation time on an object after participants had seen it for a while. It is also found that scene context have an immediate effect on object processing. In addition, Henderson and Smith (2009) investigated if visual input controls fixation duration. They conducted two scene memorization tasks and one visual search task to assess which underlying processing is related to fixation duration in the *scene onset delay* paradigm. The findings supported the mixed control model of fixation duration, implying that fixations can be under direct control. Other studies also supported the idea that fixation duration is under the direct control of the observers' scene and that fixation durations reflect moment-to-moment changes in visual and cognitive difficulty (Henderson & Pierce, 2008; Henderson et al., 2014). These studies have contributed to the idea that fixation durations can be used in measuring attention and ongoing perceptual and cognitive processes during scene viewing.

If we can use fixation duration variables for explaining cognitive processes in scene viewing, we should investigate the variables and their implication in detail. Mean fixation duration, total fixation duration (i.e., first-pass gaze duration), duration on each fixation, and the number of fixations on a particular area are some variables gathered from viewers in the eye-tracking studies (Duchowski, 2017). Previous experiments showed that the number of fixations increased as recognition performance increased (Loftus, as cited in Duchowski, 2017). In addition, the sum of all fixation duration in a region is correlated with the number of fixations in that region (Duchowski, 2017; Henderson & Hollingworth, 1998). Moreover, the average fixation duration is about 300 ms (Barthelme et al., 2013; Henderson, 2003, 2011; Nuthmann et al., 2010; Rayner, 1998), but visual and cognitive factors of the scene cause significant variability. For example, while the mean fixation duration for typing is 400ms, the mean fixation duration for silent reading is 225ms (Rayner, 1998). In addition, scenes' luminance, contrast, quality, and color, whether participant search or memorize, and amplitude of related search are some of the factors that affect fixation duration (Barthelme et al., 2013; Duchowski, 2017; Henderson, 2003; Henderson et al. 2014; Rayner, 1998). For example, in memorization tasks, individual fixation durations are longer than visual search tasks (Nuthmann et al., 2010). In addition, semantically informative (i.e., less consistent) objects cause longer first pass gaze and total fixation durations than uninformative (i.e., more consistent) objects (Barthelme et al., 2013; Duchowski, 2017; Henderson, 2011; Nuthmann et al. 2010). Moreover, the gaze control of the observer can be influenced by the object's spatial orientation (Cronin & Brockmole, 2016) and spoken language comprehension and production (Henderson, 2003). However, most of the studies on eye movements are interested in reading studies rather than gaze control on scenes (Henderson, 2011) so findings in this area are limited (Henderson, 2003; Henderson & Smith, 2009; Rayner,1998).

Studies have shown that scene context has effects on eye movement (Duchowski, 2017). Eye movements have been investigated for various information processing tasks, from art to face recognition. Early research on art and eye movements showed that information in the scene and the scene's perceptual and cognitive processing is related to the observer's eye movement patterns. For example, an examination of observers of *Sunday on the Island of La Grande-Jatte* by Georges Seurat showed that observers fixated on people rather than on the background. In addition, it is found that the complexness of the art affects the duration of eye movements (Duchowski, 2017). The difficulty of the search task also affects eye movements. Zelinsky and Sheinberg (as cited in Rayner, 1998) compared serial and parallel search tasks and showed that in the serial search task, fixations were longer, and saccades were shorter than in parallel search tasks. Memory research also used eye movement methodology. G. R. Loftus (as cited in Rayner, 1998) showed more fixations were correlated with higher recognition scores. More examples from multiple domains were given by Rayner, such as auditory language processing, problem-solving, dual tasks, face perception, and brain damage research. In auditory language processing, people's eye movements are recorded while people listen to a story, and studies showed that what people hear and where they look at is related to each other. Moreover, people tend to fixate more and longer while attending to more complicated aspects of the problem. Longer fixations also occur when people compare faces rather than looking at one face only. Therefore, it can be inferred that eye movement change according to the scene and task context.

In particular, gaze control studies conducted on visual input reveals that fixation duration is related to semantic and syntactic information processing. The most used concept to assess semantic information processing in scene viewing research has been semantic informativeness, defined as the degree to which an object in the scene is predictable. The object is less informative if the object in the scene is expected by the viewer (Henderson & Hollingworth, 1998). For example, a knife in the kitchen is less semantically informative than a copier machine in the kitchen. The interaction between semantically informative objects and eye movements has been discussed in various studies by asking viewers to find particular objects, asking them to remember the scene later, or asking how much they liked the scene (Henderson, 2017). Loftus and Mackworth (1978) found in a recognition test that viewers fixated earlier on semantically inconsistent objects, and viewers were more likely to fixate semantically informative objects after the first saccade. However, De Graef et al. (1990) contradicted the findings of Loftus and Mackworth and found no evidence. In addition, supporting De Graef et al., Henderson et al. (1999) conducted one recognition test and one visual search test by using the same paradigm and found similar findings. In addition, Võ and Henderson (2009) conducted two experiments, one of which requires later recognition of the scene and the other requires searches for target objects to show whether foveal processing of the inconsistent objects is influenced by object–scene inconsistencies. While syntactic inconsistencies are achieved by placing spatially abnormal objects, i.e., having floating objects that should rest on a surface, semantic inconsistencies are achieved by placing objects in an irrelevant place, i.e., putting a printer in the kitchen. Võ and Henderson tried to assess if they catch early eye movements in object–scene inconsistencies. Semantically inconsistent objects attracted the attention of observers resulting in more fixation count and eventually more fixation duration than semantically consistent objects. In addition, syntactically inconsistent objects resulted in the same way because of the attention necessary to resolve the object–scene inconsistency. Therefore, when objects were floating, observers fixate more and longer without considering the semantic context. In addition, observers also fixated more and longer when semantically inconsistent objects were on the surface. However, response time showed no effect on neither semantic nor syntactic inconsistency.

In addition to experiments on semantic informativeness, the relationship between fixation duration and semantic information processing is studied with computational model simulations of gaze control. Most computational models used the visual saliency hypothesis and considered where the fixations were (Henderson & Smith, 2009; Nuthmann et al., 2010). Moreover, Nuthmann et al. proposed a computational model, CRISP, that considers cognitive factors in scene viewing. The findings supported that fixation durations indicated perceptual and cognitive activity in scene viewing.

Eye movements during scene perception are conducted with different cultures and unfamiliar objects. These studies include the eye movements of English and Chinese speakers during face processing, scene perception, and visual search (Evans et al., 2009). Some studies claimed that Asian participants focus on background more than American participants because Asian culture prioritizes the group over the individual, which is the other way around in American culture. However, not all studies supported this idea. It has been shown that there were no differences between Chinese participants and American participants (Rayner et al., 2009). The mean fixation duration or proportions of fixations did not differ between Chinese and English speakers (Evens et al., 2009).

Studies with odd or emotional objects showed that participants fixated on the odd or emotional objects earlier than normal objects (Rayner et al., 2009). Merging previous studies, Rayner et al. (2009) tested Chinese and American participants with unusual/weird objects at the background to research differences how people prioritize the information to investigate the opposite opinions in cultural differences in an eye-tracking study. They found that participants showed no difference. In addition, Evans et al. (2009) studied the influence of culture during scene perception in a recognition memory test. Recognition memory and eye movement data showed that cultural differences did not affect eye movements and memory. However, both groups looked at focal objects more and longer. The claim that cultural differences affect eye movements was not supported; however, the findings of scene semantic that focal objects were looked more and longer was supported.

All in all, observers' eye movements are an essential assessment tool for perception and cognition of underlying information processing. Most of the studies in the literature have focused on reading, so there is a small number of research on scene viewing. The studies on eye movement showed that eye movements are under direct control, revealing cognitive processes (Henderson & Pierce, 2008; Henderson &

Smith, 2009; Henderson et al., 2014). In addition, there are various variables of eye movement methodology, such as fixation duration and count, and their differences are affected by the physical and semantic properties of the scene. For example, the studies have shown that the number of fixation (Henderson & Hollingworth, 1998; Henderson et al., 1999), the position of fixation (Henderson & Hollingworth, 1998; Loftus & Mackworth, 1978), and duration of fixation on a scene (Henderson, 2003) are affected by scene semantics. Moreover, several studies have shown that semantically informative objects attract longer first-pass gaze duration (i.e., total fixation duration from the first fixation to the last in a region) than uninformative objects (Henderson et al. 1999; Henderson, 2003; Henderson, 2017; Henderson & Ferreira, 2004; Võ & Henderson, 2009). In addition, it has been shown that differences in scene semantic affect the eye movements of the participants (Nuthmann et al., 2010; Rayner et al., 2009). Therefore, it can be said that perceptual and cognitive activity in scene viewing is reflected by the eye movements of observers.

In the current study, the eye movements of participants were collected while they described spatial positions of *Figure* from TRPS, which is 71 line drawings. The stimulus differences in terms of mean fixation duration, total fixation duration, and fixation count on objects were analyzed and interpreted by assuming differences in scene semantic affect the participants' eye movements, which reflect the cognitive information processing. We used previous research findings stating that eye movements reflect cognitive processes to assess the relationship between observer's spatial utterance and their eye movements during scene viewing.

Chapter 3 explains the methodology of the current study, the eye movement paradigm, the stimuli, participants, and the analysis procedure.

CHAPTER 3

METHODOLOGY

The present study aims at making an inquiry on spatial language variety in Turkish on static scenes by using the eye-tracking paradigm. The study is based on a within-subject experiment design, conducted using a Tobii T120 eye-tracker. The utterances of participants were recorded through a microphone. Moreover, participants' eye movements and response times were recorded and analyzed through eye-tracking software and statistical software JASP 0.14.0.0.

Tobii T120 is an eye tracker with 120Hz data sampling rate, typical 0.5 degrees of reported accuracy, 0.3 degrees spatial resolution, and 17" TFT screen size. Tobii T120 eye tracker communicates with the computer via a standard network cable. Participants were placed in front of the eye tracker, and the distance from the person's eyes to the eye tracker was approximately ~65 cm. The front display of Tobii T120 is presented in Figure 4 (Tobii Technology AB, 2011).

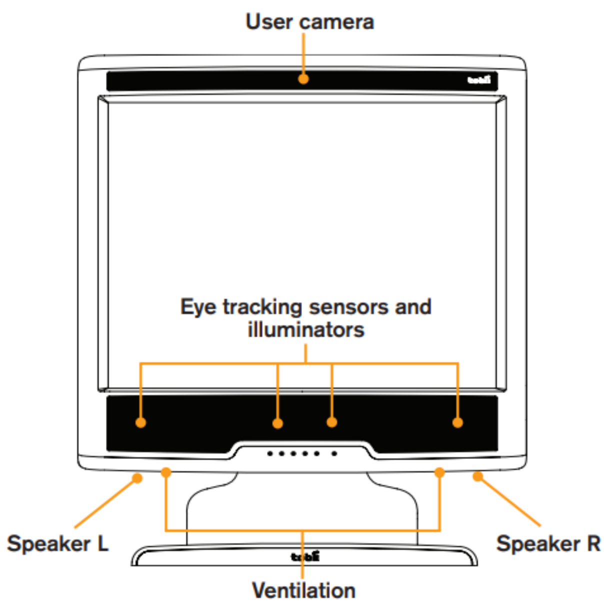

Figure 4: Front display of Tobii T120 eye tracker

According to Biederman (1987), two-dimensional drawings can be used for the recognition and naming of objects, and adults have no difficulty recognizing objects in line drawings. Moreover, scene depictions provide better control over variables in the experiments. In real-world scenes, especially in eye-tracking studies, the nature of the image, such as non-pictorial depth cues, may affect results. (Henderson, 2011). In the experiment, 71 line drawings were used as stimuli, developed by Bowerman et al. (1992). Each drawing depicted a static spatial configuration that represented a range of topological relationships. Some examples from the stimuli are shown in Figure 5.

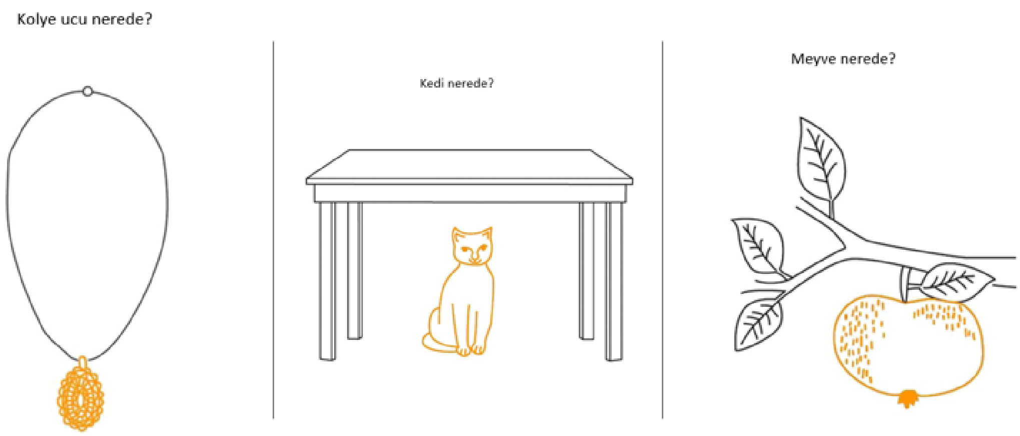

Figure 5: Some examples from the stimuli

The battery consists of scenes in which spatial terms, such as *in, on, under, over, near, against*, could be used in English. Each stimulus consists of three parts; *Figure*, *Ground*, and *Text*. In the Text part, participants were asked for each picture the question "Where is the (figure object)?" in Turkish ("(*Konumlanan* 'Figure') nerede?"). In each stimulus, the *Figure* is yellow, and the *Ground* is black.

3.1. Participants and Procedure

Thirty-four adult native adult speakers of Turkish and two adult speakers of Azerbaijani who knew Turkish participated in the experiment. The average age of participants was 21.8 years ($SD = 1.62$). All participants were presented with the same stimuli. The number of the stimuli that were presented to participants was 71. Each picture in Topological Relation Picture Series (TPRS) was presented only once.

Participants were given a form of consent, and the experimenter explained the task briefly. Participants were seated in front of the eye tracker. Eye-tracking configurations were made, including the calibration. The experiment sessions started with a page explaining the experiment to the participant, providing an example of a stimulus. The instruction page was given in APPENDIX A. Participants were presented stimuli (Topological Picture Series), and in each stimulus, participants were asked where the *Figure* was. Through the study, participants' answer was recorded through a microphone. Each participant saw the whole stimuli, including 71 pictures from TRPS. The order of the presentation of the stimuli was randomized. The participants proceeded with the stimuli in a self-paced manner without time limitations. The experiment was finalized, thanking, and asking the participant to inform the experimenter about the end of the session.

3.2. Analysis Procedure

For the analyses, the response time (RT), total fixation number on the *Figure*, the *Ground*, and the *Text* (F_{count}, G_{count}, T_{count} respectively), the average duration of one fixation on the *Figure*, the *Ground*, and the *Text* (F_{avg}, G_{avg}, T_{avg} respectively), and total duration on the *Figure*, the *Ground*, and the *Text* (F_{sum}, G_{sum}, T_{sum} respectively) were collected through eye tracker. We also recorded participants' utterances, describing where the *Figure* is in relation to the *Ground*, and divided the utterance of participants into three groups: spatial terms, reference objects, and description. Furthermore, each group was divided into *lexically rich* and *lexically poor* conditions.

We have three different research questions requiring multiple kinds of analyses. First, for our first research question, we analyzed the lexical variety of participants' utterances in terms of fixation data. Behavioral data (RT, F_{count}, G_{count}, T_{count}, F_{avg}, G_{avg}, T_{avg}, F_{sum}, G_{sum}, T_{sum}) were our dependent variables, and participants' utterances were our independent variable. Therefore, we had two conditions (i.e., *lexically rich* and *lexically poor*). The means of eye tracking variables (RT, F_{count}, G_{count}, T_{count}, F_{avg}, G_{avg}, T_{avg}, F_{sum}, G_{sum}, T_{sum}) on the *Figure* and the *Ground* for each condition (*lexically rich* and *lexically poor*) were compared through MANOVAs and ANOVAs.

Second, for our analyses of locative case markers, the independent variable was locative case markers (LOC) with two conditions: LOC rich and LOC poor. Our dependent variables were eye tracker variables (RT, F_{count}, G_{count}, T_{count}, F_{avg}, G_{avg}, T_{avg}, F_{sum}, G_{sum}, T_{sum}) and lexical richness levels of participants' utterance (lexically rich and lexically poor). Figure 6 presents two different stimuli. On the left, i.e., stimulus 34, participants were asked, *"Adam nerede?"* "Where is the man?", and on the right, i.e., stimulus 35, participants were asked, *"Yarabandı nerede?"* "Where is the band-aid?".

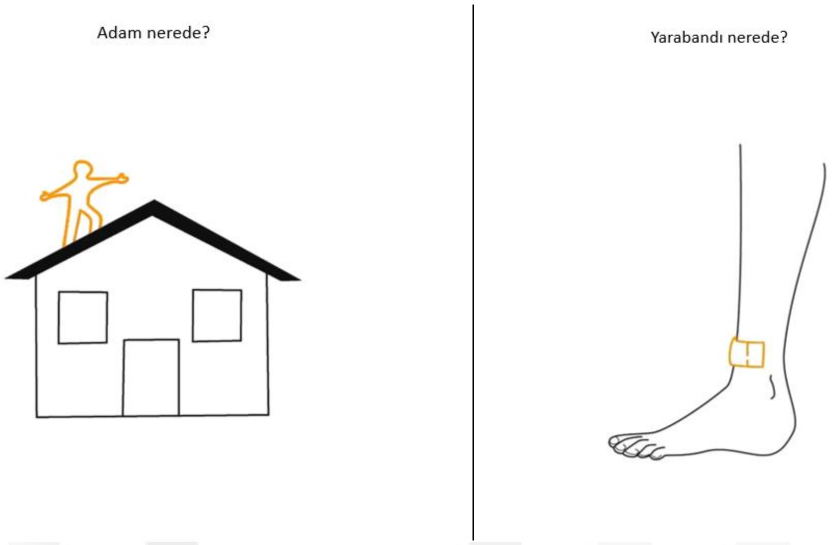

Figure 6: Two examples from the stimuli

Participants expressed the left one with two different words with locative case markers, as in (9) and (10). Therefore, the LOC variety value of stimulus 34 is two, and stimulus 34 is in the lexically poor conditions of LOC variety.

(9) *Çatıda*. ‘on the roof’

(10) E*vin çatısında*. ‘on the roof of the house’

However, participants expressed the right one with six different words with locative case markers, such as in (11), (12), and (13). Therefore, the LOC variety value of stimulus 35 is six, and stimulus 35 is in the lexically rich conditions of LOC variety.

(11) *Bacağında.* ‘on his/her leg’

(12) *Birisinin ayağında.* ‘on someone’s feet’

(13) *Bileğinde.* ‘on his/her wrist’

Third, for analyses of *Core* and *Non-Core* distinction made by Landau (2017) in Turkish, we first selected the stimuli representing *Containment* and *Support* types by using the full stimulus set provided by Landau et al. (2017). 79% (*N*= 56) of our stimuli (*N*=71) was selected to be either *Containment* or *Support* type. Then, we divided the selected stimuli into *Core* and *Non-Core* groups. Figure 7 presents examples for *Core* and *Non-Core* groups of *Support* and *Containment* types.

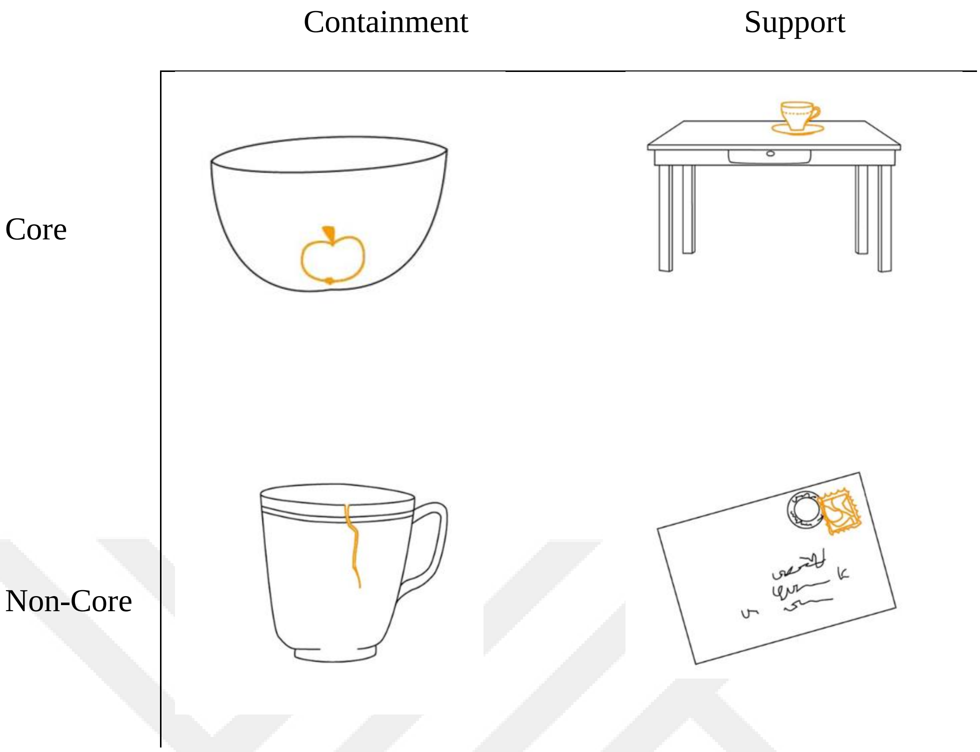

Figure 7: Examples for Core and Non-Core categorization

We conducted two different analyses. To distinguish *Core* and *Non-Core* groups in our data, we determined the independent variable as *Core* vs. *Non-Core* groups and the dependent variables as response time, fixation count on the *Figure*, the *Ground*, and the *Text*, mean fixation duration on the *Figure*, the *Ground*, and the *Text*, total fixation duration on the *Figure*, the *Ground*, and the *Text* (RT, F_{count}, G_{count}, T_{count}, F_{avg}, G_{avg}, T_{avg}, F_{sum}, G_{sum}, $T_{sum,}$ respectively) and lexical richness of participants' utterance. In addition, to investigate the differences between types of *Core* and *Non-Core* groups, we determined our independent variables as *Containment-Core*, *Containment-Non-Core*, *Support-Core*, and *Support-Non-Core* groups. Our dependent variables were eye tracker data (RT, F_{count}, G_{count}, T_{count}, F_{avg}, G_{avg}, T_{avg}, F_{sum}, G_{sum}, T_{sum}) and participants' utterances. Figure 7 presents examples of *Core* and *Non-Core* groups of *Support* and *Containment* types from our stimuli (TRPS). In Figure 7, yellow objects are the *Figure,* and the black objects are the *Ground*.

3.2.1. Utterance Analysis Procedure

Levinson et al. (2003) defined topological relation markers (TRMs) as any classes such as adpositions, spatial nouns with or without a locative case, positional verbs that code topological relations. In this study, among TRMs, we analyzed adpositions and locative case markers, and in this study, ***spatial term*** **refers to both adpositions and locative case markers.**

Participants provided 2,825 descriptions in total. The descriptions were encoded and analyzed with several working assumptions. First, in the transcription, some participants used more than one description. The first one was included, whereas the rest was excluded. Second, following the common practice in the literature, complex spatial descriptions were excluded, such as *içinden geçmiş* ‘passed through’. Third, combinations of locative case markers and a spatial term, or spatial term and visual description were used in some of the descriptions. In this case, one of them was included in the analysis. For example, in the case of *duvar-da asılı* ‘wall-LOC hanged’, LOC (wall) was included. Fourth, the ones having only visual descriptions were excluded. For instance, *masa-(y)a yapıştırılmış* ‘table-DAT sticked’ was excluded.

The last assumption was made in terms of locative case markers. In the literature on Turkish spatial terms, certain combinations of words and suffixes were included in the analysis as a locative case marker or adpositions. In addition to the adpositions that Göksel and Kerslake (2005) enlisted, the adpositions in Table 2 were analyzed, following Atak (2018).

Table 2

Some Turkish Adpositions[3]

Turkish	English
Çapraz	Cross
Köşe	Corner
Ileri	Far
Bitişik	Next to
Kenar	Side
Kısım	Part
Taraf	Side
Orta	Middle
Uç	Edge
Hiza	Line
Yakın	Close
Uzak	Far
Üzeri	Over
Teğet	Tangent

[3] The translations in the table are approximate and translated by the author.

There were spatial expressions in the current study, such as *tepesinde* ‘peak-3SG.POSS-LOC’, *başında*, ‘head-3SG.POSS-LOC’, *ağzında* ‘mouth-3SG.POSS-LOC’ that were not explicitly mentioned in the literature. To decide if these terms were the reference objects with locative case suffixes or adpositions, we employed the definition of adposition by Levinson et al. (2003). Accordingly, those words, such as *tepesinde* ‘peak-3SG.POSS-LOC’, *başında* ‘head-3SG.POSS-LOC’, *ağzında* ‘mouth-3SG.POSS-LOC’, were taken as adpositions. A supporting claim for this approach is that words for body parts describe locations in many languages. For example, in English, *foot* is used to describe mountains, e.g., “foot of the mountain” (Heine, 1997 as cited in Johanson & Papafragou, 2014). Similar words in Turkish and their English translations are presented in Table 3, which were assumed to be adpositions in the present study.

Table 3

Words Accepted as Adposition in Turkish [4]

Turkish	English
Tepesinde	On the top of
Başında	In front of/upfront
Ağzında	At the starting point
Zirvesinde	On the peak of
Dibinde	At the bottom of
Yakasında	At its side
Sırtında	On
Yamacında	Near/close to
Eteğinde	Below

[4] The translations in the table are approximate and translated by the author.

CHAPTER 4

ANALYSIS AND RESULTS

The analyses of the results included the following steps: First, we analyzed the eye fixation of participants on stimuli. Further, the relationship between eye movement data and participants' spatial expressions was inspected. To do so, we annotated participants' expressions and analyzed them according to their linguistic type. In addition, we analyzed behavioral data (i.e., eye movements and response times) according to participants' expressions.

4.1. Analyses of Gaze

In analyses of participants' fixations, there are three separate analyses of eye movements, each divides the stimuli into three parts, as presented in Table 4.

Table 4: Analysis Types

Analysis Type	**Abbreviations**	**Stimulus Area**
Mean Fixation Duration Analysis	F_{avg}	Mean Fixation Duration on the *Figure*
	G_{avg}	Mean Fixation Duration on the *Ground*
	T_{avg}	Mean Fixation Duration on the *Text*
Total Fixation Duration Analysis	F_{sum}	Total Fixation duration on *Figure*
	G_{sum}	Total Fixation duration on Ground
	T_{sum}	Total Fixation Duration on the *Text*
Fixation Count Analysis	F_{count}	Fixation count on the *Figure*
	G_{count}	Fixation count on the *Ground*
	T_{count}	Fixation count on the *Text*

Võ and Henderson (2009) define total fixation duration and total fixation count as below:

"The total fixation duration was defined as the sum of all fixation durations on the target region from scene onset until scene offset. Total fixation count was defined as the sum of all fixations located in the target region from scene onset until scene offset." (p. 6)

The following sections present mean fixation duration analysis, total fixation duration analysis, and fixation count on the *Figure* and the *Ground.*

4.1.1. Mean Fixation Duration Analysis

A repeated-measures ANOVA was conducted to determine if there were statistically significant differences in mean fixation duration between the *Figure,* the *Ground,* and the *Text,* presented in Figure 8.

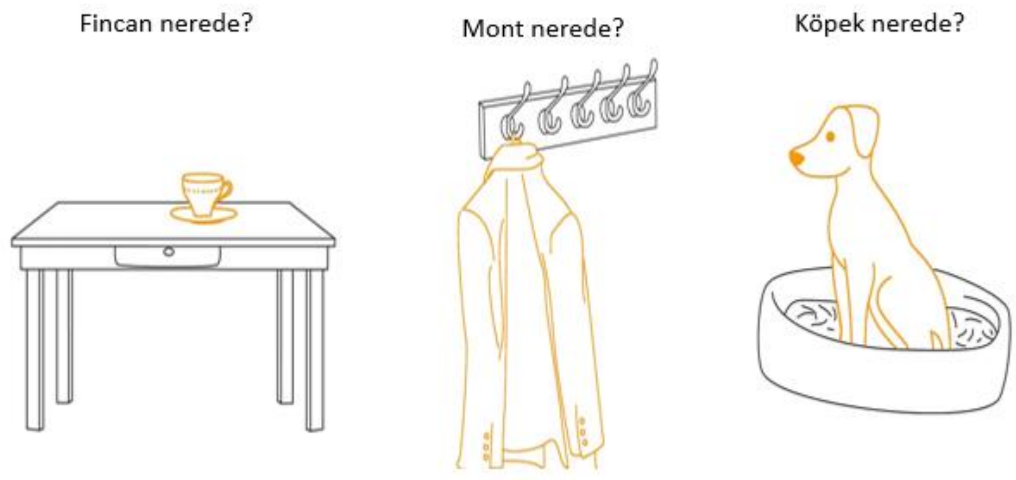

Figure 8: Stimuli examples and its parts: Figure in yellow, Ground in black, and Text

There were three outliers for mean fixation duration on the *Figure* (stimuli 7, 53, 65), as presented in Figure 9. The nature of the stimuli was considered the reason for the outlier because the *Figures* are relatively small compared to the *Grounds* in the outlier stimuli, so they were excluded from the analysis.

Table 5
Mean Fixation Duration

	Mean	SD	N
F_{avg}[5]	251	29.5	68
G_{avg}	234	23.5	68
T_{avg}	219	23.0	68

[5] F_{avg} refers mean fixation duration on the *Figure.* Similarly, G_{avg} refers mean fixation duration on the *Ground*, and T_{avg} refers mean fixation duration on the *Text.*

The data without outliers were normally distributed for each group and assessed by the Shapiro-Wilk test (*ps*=0.06, *ps*=0.24, 0.21, respectively). The assumption of sphericity was not violated, as assessed by Mauchly's Test of Sphericity, *p*=0.84. There were significant differences in mean fixation duration between the *Figure*, the *Ground*, and the *Text*, $F(2,134)=30.70$, $p<.001$, $\eta^2 = 0.31$, with higher mean fixation duration on *Figure* (*M*=251, *SD*=29.5) than mean fixation duration on *Ground* (*M*=234, *SD*=23.5), and with higher mean fixation duration on *Ground* than mean fixation duration on *Text* (*M*=219, *SD*=23.0), as presented in Table 5. Post-hoc analysis with a Holm adjustment revealed that mean fixation duration was significantly higher on the *Figure* than on the *Ground* (*M*= 17.6, $p < .001$, *d=0.52)* and on the *Text* (*M*=32.1, $p <.001$, *d*=0.95*)*. Moreover, mean fixation duration was significantly higher on *Ground* than on *Text* (*M*=14.5, $p < .001$, *d*=0.43.)

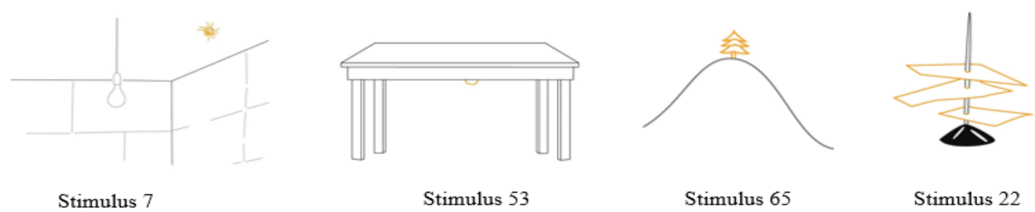

Figure 9: Excluded outliers from the analyses

In sum, the stimuli were divided into the *Figure*, the *Ground*, and the *Text*. The analysis of mean fixation duration revealed that participants fixated more on each fixation on the *Figure* than the *Ground*, and they also fixated more on each fixation on the *Ground* than the *Text*. In the following section, we analyzed the total fixation duration on stimuli.

4.1.2. Total Fixation Duration Analysis

Total fixation duration is the variable showing the sum of all fixation duration on a specified part of the stimuli, such as on the *Figure*, the *Ground*, or the *Text*. A repeated-measures ANOVA was conducted to determine if there were statistically significant differences in total fixation duration between the *Figure*, the *Ground*, and the *Text*. There were three outliers for total fixation duration on the *Text* and two outliers for total fixation duration on the *Figure*. One outlier (stimulus 22, as presented in Figure 6[6]) has several *Figures*, so we excluded it from the total fixation duration analysis. The data without the excluded outlier was normally distributed for the *Figure* and the *Ground* group but not for

[6] Stimulus 7, 53, and 65 were excluded from mean fixation duration analysis, while Stimulus 22 was excluded from total fixation duration and fixation count analyses.

the *Text* group, assessed by the Shapiro-Wilk test (*ps*=0.53, *ps*=0.18, *ps*= < .001, respectively). The assumption of sphericity was violated, as assessed by Mauchly's Test of Sphericity, *p*= <.001. Therefore, a Greenhouse-Geisser correction was applied (ε = 0.75).

There are significant differences in total fixation duration between the *Figure*, the *Ground*, and the *Text*, $F(1.5,103)=113$, $p<.001$, $\eta^2 = 0.62$, with higher mean fixation duration on the *Figure* (*M*=32502 *SD*=12082) than mean fixation duration on the *Ground* (*M*=23996, *SD*=12384), and with higher mean fixation duration on *Ground* than mean fixation duration on the *Text* (*M*=6892, *SD*=2512), as presented in Table 6. Post-hoc analysis with a Holm adjustment revealed that mean fixation duration was significantly higher on the *Figure* than on the *Ground* (*M*= 8505, *p* < .001, *d*=0.58), and on the *Text* (*M*=25610, *p* <.001, *d*=1.76). Moreover, mean fixation duration was significantly higher on the *Ground* than on the *Text* (*M*=17104, *p* < .001, *d*=1.17).

Table 6
Total Fixation Durations[7]

	Mean	SD	N
F_{sum}	32502	12082	68
G_{sum}	23996	12384	68
T_{sum}	6892	2512	68

In summary, the analysis of total fixation duration revealed that participants fixated more in total on the *Figure* than the *Ground*. They also fixated more in total on the *Ground* than the *Text*. In the following section, we analyzed the fixation counts on stimuli.

4.1.3. Fixation Count Analysis

Fixation count is a variable that presents the sum of fixations on a specified part of the stimuli, such as on the *Figure*, the *Ground*, or the *Text*. A repeated-measures ANOVA was conducted to determine if there were statistically significant differences in fixation count between the *Figure*, the *Ground*, and the *Text*. There were some outliers in the data. One outlier (stimulus 22) in the *Figure* data has several *Figures*, so it was excluded from the fixation count analysis. The data without the excluded outlier was normally distributed for the *Figure* and the *Ground* group but not for the *Text* group, assessed by the Shapiro-Wilk test (*ps*=0.11, *ps*=0.40, *ps*= < .001, respectively). The assumption of sphericity was violated, as assessed by Mauchly's Test of Sphericity, p= <.001. Therefore, a Greenhouse-

[7] F_{sum} means total fixation duration on the *Figure*. Similarly, G_{sum} means total fixation duration on the *Ground*, and T_{sum} means mean fixation duration on the *Text*.

Geisser correction was applied (ε = 0.75). There were significant differences in fixation count between the *Figure*, the *Ground*, and the *Text*, $F(1.5,104)=96.1$, $p<.001$, $\eta^2 = 0.58$, with higher fixation count on the *Figure* (*M*=129, *SD*=51.0) than fixation count on *Ground* (*M*=102, *SD*=51.8), and with higher fixation count on the *Ground* than fixation count on *Text* (*M*=31.2, *SD*=10.4), as listed in Table 7. Post-hoc analysis with a Holm adjustment revealed that fixation count was significantly higher on the *Figure* than on the *Ground* (*M*= 26.5, $p < .001$, *d*=0.43) and on the *Text* (*M*=97.8, $p <.001$, *d*=1.60). Moreover, fixation count was significantly higher on *Ground* than on *Text* (*M*=71.371, $p < .001$, *d*=1.16).

Table 7
Fixation Counts[8]

	Mean	SD	N
F_{count}	129	51.0	70
G_{count}	102	51.8	70
T_{count}	31.2	10.4	70

Overall, participants fixated the *Figure* more and longer (mean/total) than the *Ground*, and they also fixated *Ground* more and longer (mean/total) than the *Text*. To be able to understand these differences, we further investigate participants' eye movements in relation to participants' utterances. The following section presents the results of analyses of the lexical variety of participants' utterances.

4.2. Analyses of Lexical Variety of Participants' Utterance

In total, participants produced 2,528 utterances. The mean number of words used for expressing each stimulus (TRPS) was 2.05 (*SD*=0.34, *Range* = 1.33-3.19).

Approximately 10% (*N* = 224) included expressions that cannot be labeled as static spatial expressions. Accordingly, a total of 2,304 utterances were analyzed, and the analyses are reported in this section (91% of all the data points).

Approximately one-third (*N* = 817) of the utterances included locative case markers suffixed to words, indicating that locative case markers had major importance in Turkish

[8] F_{count} means fixation count on the *Figure.* Similarly, G_{count} means fixation count on the *Ground*, and T_{count} means fixation count on the *Text.*

static spatial language. Therefore, a separate analysis of locative case markers is essential for a better understanding of the topic.

The remaining utterances, which did not include locatives (N = 1,487 of 2,304), included adpositions.

In terms of frame of reference (FoR), participants used relative FoR and intrinsic FoR. Out of 2,528 utterances, relative FoR was used only for 3% (N=76). Spatial expressions, including locative case markers, were considered to be in the intrinsic FoR group. Participants used intrinsic FoR was for 88% ($N=2228$) of the utterances. 56% of them included adposition, and 32% included LOC.

In light of variables of eye-tracking data, we examine the data in accordance with the research questions. The first research question of the study was "What is the relationship between oculomotor and behavioral variables (e.g., fixation duration, fixation count, response time) and stimuli components (i.e., the *Figure* and the *Ground*) in relation to the variety of spatial expressions in Turkish?"

Descriptions of the participants as in (14), in the current study, consisted of two parts: **spatial terms (adpositions and LOCs)** as in (15), and **reference objects** as in (16).

(14) *Masanın altında.* 'under the desk'

(15) *altında* 'under'

(16) *masanın* 'desk'

Participants used a variety of words to express spatial terms and reference objects. To describe the relationship between the arrow and the apple in Figure 10, each participant produced one spatial expression.

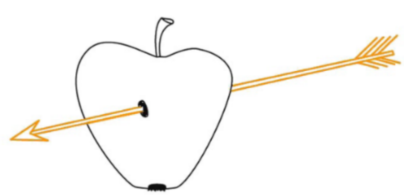

Figure 10: Stimulus 30

Out of 36 descriptions that 36 participants produced, we had four different descriptions as in (17), two different reference objects as in (18), and three different spatial terms as in (19).

(17) *Elmanın içinde*. 'in the apple'

Elmanın ortasında. 'in the middle of the apple'

Elmanın içerisinde. 'inside the apple'

Meyvenin içinde 'in the fruit'

(18) *elmanın* 'apple'

meyvenin 'fruit'

(19) *içinde* 'in'

ortasında 'in the middle of'

içerisinde 'inside'

Participants' utterances vary to express a spatial configuration in Turkish, as exemplified in (17),(18), and (19). Therefore, we analyzed the difference between the variation of spatial terms and reference objects. A repeated-measures ANOVA was conducted to determine if there were statistically significant differences in the number of different words used for spatial terms (locative case markers included) and reference objects. There were outliers, occupying 6% of data (stimulus 17 and stimulus 46 for spatial terms, and stimulus 47 and stimulus 18 for reference objects), mostly due to large differences in the representation of the *Figure* and the *Ground* in the stimuli. On the other hand, they were not excluded from the analysis since they did not have a significant impact on statistical results.

The results revealed that there were significant differences in the number of different words used for spatial terms and reference objects, $F(1,70)= 14.4$, $p <.001$, $\eta^2 = 0.17$, with the number of different words for spatial terms (M=6.50, SD=3.58) being less than reference objects (M=4.73, SD=3.22). Post-hoc analysis with a Holm adjustment revealed that the number of different words used was significantly decreased from spatial terms to reference objects (M= 1.77, $p < .001$, d=0.45). More specifically, the number of the different spatial terms and words, including LOC, was statistically higher than the number of varying reference objects used for *Ground*.

Therefore, we have conducted further analyses by dividing the utterances into three groups, namely spatial terms, reference objects, and their combination (viz. descriptions), for a detailed assessment of their variety in the utterances.

1. Spatial Terms Variety Analysis: The analysis of different spatial terms and locative case markers used to describe the position of the *Figure* in accordance to the *Ground* for each stimulus.

2. Reference Object Variety Analysis: The analysis of different lexemes used to describe the *Ground* for each stimulus.

3. Description Variety Analysis: The analysis of different spatial utterance participants made i.e., combinations of spatial terms and reference objects in each stimulus.

In the data, the same descriptions were used by different participants for different stimuli. There were both unique and repeated expressions in the data. Out of 2,304 statical spatial descriptions, approximately 26% (*N*=616) included different static spatial descriptions. Approximately 19% (*N*=438) were found in the data just once. The rest, 7% (*N*=178), were produced repeatedly by participants for different stimuli. Similarly, the number of different spatial terms covers approximately 9% (*N*=204) of the data. While adpositions cover 3% (*N*=60), locative case markers cover 6% (*N*=144). In addition, 9% (*N*=201) of the data include different reference objects. These numbers are the variety of spatial expressions in the participants' utterances.

In our analyses, we calculated the variety of spatial terms, reference objects, and descriptions for each stimulus. For example, for stimulus 53, all participants produced *Masanın altında* "under the desk" so the value of description variety of stimulus 53 is one. However, for stimulus 30, as mentioned in examples (17), (18), and (19), *Elmanın içinde* "in the apple", *Elmanın ortasında* "in the middle of the apple", *Elmanın içerisinde* "inside the apple", and *Meyvenin içinde* "in the fruit" were produced by participants. Therefore, the value of description variety of stimulus 30 is four. We calculated the mean values of spatial terms variety, reference objects variety, and description variety. In the data, the mean of description variety for each stimulus was 10,1. The descriptions are combinations of spatial terms and reference objects. The mean of spatial terms variety for each stimulus was 6,50, and the mean of reference object variety for each stimulus was 4,70. Later, we labeled stimuli with a variety value less than mean as lexically poor. Furthermore, we labeled stimuli with a variety value more than mean as lexically rich.

In summary, descriptions were divided into two: spatial terms and reference objects. In the current study, Turkish spatial terms are made of LOCs and adpositions. A total of 2,304 utterances were analyzed, and approximately one-third of the utterances included locative case markers. The rest included adpositions. Some adpositions and words with LOCs were used repeatedly by participants. We determined unique terms for each stimulus and analyzed the variety of spatial terms, reference objects, and adpositions. The following section investigated the relationship between participants' gaze and the variety of participants' spatial expressions.

4.3 The relationship between Gaze Data and Spatial Descriptions

In our study, we asked if there was a relationship between fixation data and the variety of spatial descriptions. Spatial terms variety, reference objects variety, and description variety were analyzed in terms of fixation data on the *Figure* and the *Ground*. Fixation data has three variables: mean fixation duration, total fixation duration, fixation count. Mean fixation duration on the *Figure* (F_{avg}), total fixation duration on the *Figure* (F_{sum}), total fixation count on the *Figure* (F_{count}), mean fixation duration on the *Ground* (G_{avg}), total fixation duration on the *Ground* (G_{sum}), and total fixation count on the *Ground* (G_{count}) are the dependent variable. We start analyses with spatial terms and continue with reference objects, and finish with descriptions. For example, Figure 11 represents a spatial relationship between a cup and a desk. First, we divide the unique description of participants for the stimulus, as in (20), into two: spatial terms as in (21) and reference objects as in (22).

Figure 11: Stimulus 1

(20) *Masanın üzerinde.* 'on the table'

Masanın üstünde. 'on the table'

Masanın orta sağında. 'in middle right of the table'

Masanın sağ üst köşesinde. 'in the upper right corner of the table'

Kasesinin üzerinde 'on its bowl(coaster)'

(21) *üzeride* 'on'

üstünde 'on'

orta sağında 'in middle right of'

sağ üst köşesinde ‘in the upper right corner of’

(22) *masa* ‘the desk’

kase ‘the bowl(coaster)’

We calculated the value of spatial terms variety, reference objects variety, and descriptions variety. In this example, the description variety value is five. The spatial terms variety value is four, and reference object variety is two. Then we labeled *lexically poor and lexically rich* conditions to stimuli according to the values of spatial terms variety, reference objects variety, and description variety. Therefore, we had two conditions (i.e., *lexically rich* and *lexically poor*) for eye tracking variables. The means of eye tracking variables (RT, F_{count}, G_{count}, F_{avg}, G_{avg}, F_{sum}, G_{sum}) on the *Figure* and the *Ground* for reach conditions (*lexically rich* and *lexically poor*) were compared through MANOVAs and ANOVAs.

The analyses in the following sections showed that, in general, speakers fixated on the *Figure* and the *Ground* more and longer when they produced richer spatial terms, reference objects, and descriptions. In other words, topological relations, inducing lexically rich descriptions, also induce longer fixation duration (mean and total) and fixation count on stimuli. The average number of descriptions did not change significantly between stimuli, so the reason for differences of fixation count and duration on stimuli were not caused by the number of the descriptions that participants produced; rather, it was the lexical variety that participants produced. When the lexical variety of populations’ utterance was higher, fixation duration (mean and total) and fixation count on stimuli were higher.

4.3.1. Spatial Terms Variety

To answer the question “What is the relationship between oculomotor and behavioral variables (e.g., fixation duration, fixation count, response time) in relation to the variety of spatial terms in Turkish?”, we conducted three MANOVA and several one-way ANOVA tests to determine if the mean and total fixation duration, and fixation count on the *Figure* and the *Ground* (F_{avg}, G_{avg}, F_{sum}, G_{sum}, F_{count}, F_{count}) were different for groups with different lexical richness in spatial terms.

The mean number of different spatial terms used by participants in each stimulus was calculated. Stimuli were classified into two groups: lexically rich (*N*= 28) and lexically poor (*N*=43). The ones less than mean were labeled as lexically poor, whereas those more than mean were labeled as lexically rich. Figure 12 presents stimulus 16 on the left and stimulus 17 on the right.

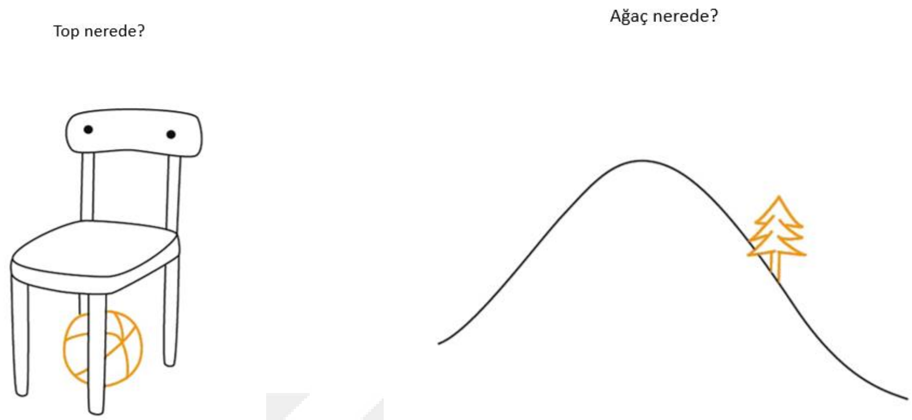

Figure 12: Lexically rich and lexically poor conditions for spatial terms

Participants produced 19 different spatial terms for stimulus 17, depicting a tree in a mountain. Nevertheless, participants produced two different spatial terms for stimulus 16, depicting a ball under the chair. While stimulus 17 was labeled as rich, stimulus 16 was labeled as poor. We started with reporting differences of mean fixation duration, then continued total fixation duration, and finally fixation count on the *Figure* and the *Ground* between lexically rich and lexically poor groups.

Firstly, a one-way multivariate analysis of variance (MANOVA) was conducted to test the hypothesis that there would be differences between mean fixation durations on the *Figure* and the *Ground* between lexical richness levels. Two outliers, stimulus seven and stimulus 22, approximately 3% of the data, were excluded from the mean fixation duration analysis. Using Pillai's trace, there was a significant effect of spatial terms variety on the mean fixation duration on the *Figure* and the *Ground* (F_{avg} and G_{avg}), $V = 0.13$, $F(2, 66) = 5$, $p = 0.01$. Prior to conducting a series of follow-up ANOVAs, the homogeneity of variance assumption was tested. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.30). Two one-way ANOVAs on dependent variables (F_{avg} and G_{avg}) were conducted as follow-up tests to the MANOVA. Both ANOVAs were statistically significant.

One-way ANOVA test revealed that mean fixation duration on the *Figure* (F_{avg}) was significantly different between lexically rich and lexically poor conditions of spatial terms, $F(1, 67) = 5.67$, $p = .02$, $\eta^2 = 0.08$. Two outliers, as assessed by boxplots, stimuli seven from the rich condition and stimuli 53 from the poor condition, were excluded from the analysis. The data without outliers were normally distributed for each group and assessed by the Shapiro-Wilk test (*ps*=0.26, *ps*=0.23). There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p =0.14). The fixation duration mean was lower in poor condition (M= 246, SD= 26.6) than the rich condition (M= 264, *SD*= 35), as listed in Table 8. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p = .02, d= -0.59).

Table 8
Mean Fixation Durations on Figure

Spatial Terms Richness	Mean	SD	N
Poor	246	26.6	42
Rich	264	35.0	27

In addition, the one-way ANOVA test showed that mean fixation duration on the *Ground* (G_{avg}) was statistically significantly different between lexically rich and lexically poor conditions in spatial terms, $F(1, 69) = 6.57$, p =.013, $\eta^2 = 0.087$. There were no outliers, as assessed by boxplots. Each group was normally distributed, assessed by the Shapiro-Wilk test. As assessed by Levene's test of homogeneity of variances (p= 0.65), there was homogeneity of variances for G_{avg}. The fixation duration mean was lower in poor condition (*M*= 229, *SD*= 21.8) than the rich condition (*M*= 244, *SD*= 25.6), presented in Table 9. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p = .0013, d= -0.62).

Table 9
Mean Fixation Durations on Ground

Spatial Terms Richness	Mean	SD	N
Poor	229	21.8	43
Rich	244	25.6	28

Secondly, we analyzed the total fixation duration on the *Figure* and the *Ground*. Differences of total fixation duration on the *Figure* (F_{sum}) and the *Ground* (G_{sum}) between lexically rich and lexically poor conditions were analyzed with one-way MANOVA. Stimulus 22 was excluded from data as an outlier assessed by box plots. A statistically significant MANOVA effect was obtained, *Pillais' Trace* = 0.20, $F(2, 67) = 8.60$, $p <$ .001. Prior to conducting a series of follow-up ANOVAs, the homogeneity of variance assumption was tested. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.35) However, separate univariate tests on the outcome variables revealed that total fixation duration on *Figure* (F_{sum}) was not statistically significantly different between lexically rich and lexically poor conditions of spatial terms.

One-way ANOVA was conducted to assess if there were differences in total fixation duration on the *Ground* between poor and rich conditions. Total fixation duration on *Ground* (G_{sum}) was statistically significantly different between lexically rich and lexically poor conditions for spatial terms, $F(1, 69) = 17.0$, $p < .001$, η^2 = 020. There were no outliers, as assessed by boxplots. Each group was normally distributed, assessed by the Shapiro-Wilk test. As assessed by Levene's test of homogeneity of variances (*p*=0.21), there was homogeneity of variances for G_{sum}. The fixation duration mean was lower in poor condition (*M*= 19715, *SD*= 10649) than the rich condition (*M*= 30861, *SD*= 11828), as presented in Table 10. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant ($p < .001$, *d*= -1.0).

Table 10
Total Fixation Durations on Ground

Spatial Terms Richness	Mean	SD	N
Poor	19715	10649	43
Rich	30861	11828	28

Finally, differences of fixation count on *Figure* (F_{count}) and *Ground* (G_{count}) between lexically rich and lexically poor conditions were analyzed with one-way MANOVA. A statistically significant MANOVA effect was obtained, *Pillais' Trace* = .16, $F(2, 68) = 6.33$, $p = .003$. Prior to conducting a series of follow-up ANOVAs, the homogeneity of variance assumption was tested. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (*p*=0.82). However, separate univariate tests on the outcome variables revealed that fixation count on the *Figure* (F_{count}) was not statistically significantly different between lexically rich and lexically poor conditions of spatial terms.

One-way ANOVA was conducted to assess if there were differences of fixation count on the *Ground* between poor and rich conditions. Total fixation count on the *Ground* (G_{count}) was statistically significantly different between lexical rich and lexical poor conditions for spatial terms, $F(1, 69) = 12.5$, $p < .001$, $\eta^2 = 0.15$. There were no outliers, as assessed by boxplots. Each group was normally distributed, assessed by the Shapiro-Wilk test. As assessed by Levene's test of homogeneity of variances (*p*=0.51), there was homogeneity of variances for G_{sum}. Fixation count was higher in the rich condition (*M*= 127, *SD*= 49.0) than the poor condition (*M*= 86.5, *SD*= 46.8), as presented in Table 11. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant ($p < .001$, *d*= -0.86)

Table 11
Fixation Count on Ground

Spatial Terms Richness	Mean	SD	N
Poor	86.5	46.8	43
Rich	127	49.0	28

From the statistical analysis above, it can be interpreted that participants fixated more, spent more time looking in total, and spent more time looking at one fixation at *Ground* in stimulus, in which participants used richer vocabulary to express topological relationships between the *Figure* and the *Ground*, i.e., spatial terms. In addition, there were no differences between lexically rich and lexically poor conditions on fixation count and total fixation duration on the *Figure*. However, participant's mean fixation duration was longer on *Figure* for stimuli in which participants produced lexical rich spatial expressions. In the following section, similar to spatial term variety, the relationship between participants' gaze and reference objects variety was investigated.

4.3.2. Reference Object Variety

Descriptions consist of spatial terms and reference objects. For example, when participants produced *masanın üstünde* 'on the table', *masa* 'table' is a reference object and *üstünde* 'on' is a spatial term. Participants used a variety of different reference objects during the task. To answer the research question if the variety of reference object terms in Turkish produce any differences in fixation duration and count on the *Figure* and the *Ground*, we analyzed reference objects.

Three one-way MANOVA and several one-way ANOVA tests were conducted to determine if the fixation count, mean and total fixation duration on *Figure* (F_{count}, F_{avg}, F_{sum}) and fixation count, mean and total fixation duration on *Ground* (G_{count}, G_{avg}, G_{sum}) were different for conditions with different lexical richness levels in reference object. Stimuli were classified into two conditions: lexically rich (N=35), lexically poor (N=36) in terms of reference objects. Each group, except F_{avg}, was normally distributed, assessed by the Shapiro-Wilk test. P values of Shapiro-Wilk were 0.014 and 0.002 for poor and rich conditions of F_{avg}, respectively. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p= 0.33,0.88,0.54,0.12,0.20,0.15), for F_{sum}, F_{count}, F_{avg}, G_{sum}, G_{count}, and G_{avg}, respectively. We divided reference object variety analysis into three parts: mean fixation duration on *Figure* and *Ground*, total fixation duration on *Figure* and *Ground*, fixation count on *Figure* and *Ground*.

Firstly, a one-way multivariate analysis of variance (MANOVA) was conducted to if there were mean differences of mean fixation durations on the *Figure* and the *Ground* (F_{avg} and G_{avg}) between lexical richness levels of reference objects. Two outliers, stimulus seven and stimulus 22, were excluded from the data. Using Pillai's trace, we found no significant effect of reference objects variety on the mean fixation duration on the *Figure* and the *Ground* (F_{avg} and G_{avg}). Two one-way ANOVAs on dependent variables (F_{avg} and G_{avg}) were conducted as follow-up tests to the MANOVA. Both ANOVAs were insignificant.

Secondly, a one-way multivariate analysis of variance (MANOVA) was conducted to test the hypothesis that there would be significant differences between total fixation durations on the *Figure* and the *Ground* between lexical richness levels. Using Pillai's trace, there was a significant effect of spatial terms variety on the total fixation duration on *Figure* and *Ground* (F_{sum} and G_{sum}), $V = 0.20$, $F(2, 68) = 8.78$, $p < 0.001$. Prior to conducting a series of follow-up ANOVAs, the homogeneity of variance assumption was tested. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.07). Two one-way ANOVAs on dependent variables (F_{sum} and G_{sum}) were conducted as follow-up tests to the MANOVA. Both ANOVAs were statistically significant.

Follow-up one-way ANOVA test revealed that total fixation duration on *Figure* (F_{sum}) was statistically significantly different between lexical rich and lexical poor conditions for reference terms, $F(1, 69) = 5.60$, p =0.02, $\eta^2 = 0.08$. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p= 0.34). The fixation duration mean was higher on the rich condition (M= 36916, SD= 14105) than the poor condition (M= 29501, SD= 11291), as presented in Table 12. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p = .002, d= -0.58).

Table 12
Total Fixation Duration on Figure

Reference Object Richness	Mean	SD	N
Poor	29501	11291	36
Rich	36916	14105	35

Moreover, the one-way ANOVA test revealed that total fixation duration on the *Ground* (G_{sum}) was statistically significantly different between lexically rich and lexically poor conditions for reference terms, $F(1, 69) = 5.60$, p =0.004, $\eta^2 = 0.11$. Stimulus 66, in poor condition, was one outliner assessed by boxplots. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p= 0.13). The fixation duration mean was higher on the rich condition (M= 28257 SD= 12711) than the poor condition (M= 20080, SD=10646), as presented in Table 13. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p = .003, d= -0.54).

Table 13
Total Fixation Duration on Ground

Reference Object Richness	Mean	SD	N
Poor	20080	10646	36
Rich	28257	12711	35

Finally, mean differences between fixation count on *Figure* and *Ground* in terms of lexical richness levels of reference objects were tested with a one-way multivariate analysis of variance (MANOVA). There was a significant effect of reference object variety on the fixation count on *Figure* and *Ground* (F_{count} and G_{count}), $V = 0.18$, $F(2, 67) = 7.50$, p =

0.001. Prior to conducting a series of follow-up ANOVAs, the homogeneity of variance assumption was tested. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.32). Two one-way ANOVAs on dependent variables (F_{count} and G_{count}) were conducted as follow-up tests to the MANOVA. Both ANOVAs were statistically significant.

Fixation count on the *Figure* (F_{count}) was statistically significantly different between lexical rich and lexical poor conditions for reference terms, $F(1, 69) = 5.60$, p =0.03, η^2 = 0.07. One outlier, stimulus 22, was excluded from the analysis. The fixation count mean was higher in the rich condition (*M*= 143, *SD*= 50.1) than the poor condition (*M*=116, *SD*=49.0), as listed in Table 14. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p = .003, d= -0.54).

Table 14
Fixation Count on Figure

Reference Object Richness	Mean	SD	N
Poor	116	49.0	36
Rich	143	50.1	34

In addition, a significant one-way ANOVA effect was obtained, $F(1, 69) = 7.10$, p =0.01, η^2 = 0.09 for fixation count on the *Ground* between lexical richness conditions. Stimulus 66, in poor condition, was one outliner, assessed by boxplots. The mean of fixation count increased was higher for the rich condition (*M*= 119, *SD*= 53.3) than the poor condition (*M*= 87.2, *SD*=45.2), as presented in Table 15. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p = .01, d = -0.63).

Table 15
Fixation Count on Ground

Reference Object Richness	Mean	SD	N
Poor	87.2	45.2	36
Rich	119	53.3	34

In sum, the statistical results show that participants fixated more and spent more time looking at the *Figure* and the *Ground* for stimulus in which participants used richer vocabulary to express reference objects. However, their mean fixation duration on the *Figure* and the *Ground* were not different between lexical rich and lexical poor conditions. In the following section, we investigated the relationship between participants' gaze and descriptions, which are a combinations of spatial terms and reference objects.

4.3.3.Description Variety

In the analysis before, the variety of spatial terms and reference objects was discussed. In this part, description variety, the combination of spatial terms and reference objects, is discussed. For example, participants produced three different descriptions for stimulus 2, depicting an apple in a bow, as in (23), (24), and (25)

(23) *Tabağın içinde.* 'in/inside the plate'

(24) *Bardağın içinde.* 'in/inside the glass'

(25) *Tabağın en altında.* 'at the bottom of the plate'

For this stimulus, the variety of spatial terms is two; *içinde* 'in/inside', and *en altında* 'at the bottom'. The variety of reference objects is two; *tabak* 'plate' and *bardak* 'glass'. In addition, the variety of descriptions is three, all of the descriptions.

Similar to the analysis of spatial terms and reference objects, the mean value of descriptions variety was calculated. The stimuli with a lower mean were labeled as lexically poor, and the stimuli with a higher mean were labeled as lexically rich. Fixation count and mean, and total duration of fixation on the *Figure* and the *Ground*, i.e., F_{count}, G_{count}, F_{avg}, G_{avg}, F_{sum}, and G_{sum}, between description variety conditions were analyzed by using three distinct MANOVAs and several follow-up ANOVAs. The data was not normally distributed for the poor and rich groups of F_{avg}, and the rich group of F_{count}, assessed by the Shapiro-Wilk test (*ps*<0.001, *ps*=0.04, *ps*=0.03). The rest was normally distributed.

Firstly, a one-way multivariate analysis of variance (MANOVA) was conducted to test if there were significant differences between mean fixation durations on the *Figure*, fixation durations on the *Ground*, between the levels of descriptions variety. Three outliers, stimulus seven, stimulus 53, and stimulus 65, were excluded from the data. A statistically significant MANOVA effect was obtained on the mean fixation duration on *Figure* and *Ground* (F_{avg} and G_{avg}), $V = 0.11$, $F(2, 65) = 4.10$, $p = 0.02$. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.46). Two one-way ANOVAs on dependent variables (F_{avg} and G_{avg}) were conducted as follow-up tests to the MANOVA. However, there were no significant differences between independent variables on F_{avg}.

One-way ANOVAs were conducted to determine if the fixation duration means on *Ground* (G_{avg}) was different for groups with poor or rich levels of descriptions. There was one outlier in the rich condition, stimuli 22, as assessed by boxplots. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.23). The fixation duration mean on *Ground* (G_{avg}) were statistically significantly different between poor and rich groups $F(1, 69) = 6.62$, $p = 0.01$, $\eta^2 = 0.09$.

The mean fixation duration on the *Ground* was lower in the poor group (*M*= 228, *SD*=24.3) than the rich group (*M*= 243, *SD*= 22.2), as presented in Table 16. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p =0.01, d= -2.57). Therefore, the mean of participants' fixation duration on the *Ground* was more when their description variety was more.

Table 16
Mean Fixation Durations on Ground

Decription Richness	Mean	SD	N
Poor	228	24.3	37
Rich	243	22.2	34

Second, differences between means of total fixation durations on *Figure* (F_{sum}) and *Ground* (G_{sum}) in terms of description variety was assessed with multivariance one-way ANOVA (one-way MANOVA). One outliner, stimulus 22, was excluded from the data. There was a significant effect of reference object variety on the fixation count on the *Figure* and the *Ground* (F_{sum} and G_{sum}), V = 0.29, $F(2, 67) = 13.4$, $p < .001$. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.08). Two one-way ANOVAs on dependent variables (F_{sum} and G_{sum}) were conducted as follow-up tests to the MANOVA. Both ANOVAs were statistically significant.

One-way ANOVA was conducted to determine if the total fixation durations on the *Figure* (F_{sum}) were different for groups with poor or rich levels of descriptions. There was one outlier in each group, stimuli 22 and stimuli 25, in the poor and rich conditions, as assessed by boxplots. Variances were homogeneous, as assessed by Levene's test of homogeneity of variances (p= 0.87). The total fixation duration on the *Figure* (F_{sum}) was statistically significantly different between poor and rich conditions, $F(1, 69)$ =6.83, p = 0.01, η^2 = 0.09. The mean of total fixation duration on the *Figure* was lower for the poor group (*M*= 29384, *SD*= 12180) than the rich group (*M*= 37261.09, *SD*= 13212.34), as presented in Table 17. Tukey post hoc analysis revealed that the mean decrease from poor to rich was statistically significant (p =.011, d= -2.61). Therefore, participants' total fixation duration on the *Figure* was longer when their descriptions vary more.

Table 17
Total Fixation Durations on Figure

Decription Richness	Mean	SD	N
Poor	29384	12180	37
Rich	37261	13212	34

In addition, one-way ANOVA was conducted to determine if the total fixation durations on the *Ground* (G_{sum}) were different for groups with poor or rich levels of descriptions. There were two outliers, stimuli 63 and 24, in the poor group of G_{sum}, as assessed by boxplots. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.42). The total fixation duration on the *Ground* (G_{sum}) was statistically significantly different between poor and rich conditions, $F(1, 69) = 16.45$, $p < .001$, $\eta^2 = 0.19$. The mean of the total fixation duration on the *Ground* was lower in the poor group (*M*= 18960, *SD*=10840) than the rich group (*M*= 29718, *SD*= 11503), as presented in Table 18. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant ($p < .001$, *d*= -4.06). Therefore, participants' fixation duration on the *Ground* was more when their description variety was more.

Table 18
Total Fixation Durations on Ground

Decription Richness	Mean	SD	N
Poor	18960	10840	37
Rich	29718	11503	34

Lastly, a multivariance ANOVA test was used to understand the differences between fixation count on the *Figure* (F_{count}) and the *Ground* (G_{count}) in terms of description variety. Using Pillai's trace, there was a significant effect of description variety on the fixation count on the *Figure* and the *Ground* (F_{count} and G_{count}), $V = 0.24$, $F(2, 68) = 10.6$, $p < 0.001$. Prior to conducting a series of follow-up ANOVAs, the homogeneity of variance assumption was tested. Based on Box's Test for Equivalence of Covariance Matrices, the homogeneity of variance assumption was considered satisfied (p=0.25). Two follow-up one-way ANOVAs on dependent variables (F_{count} and G_{count}) were statistically significant.

One-way ANOVA was conducted to determine if the fixation counts on the *Figure* (F_{count}) were different for groups with poor or rich levels of descriptions. There were two outliers, stimuli 22 and 70, in the rich condition, as assessed by boxplots. Variances were homogeneous, as assessed by Levene's test of homogeneity of variances (p= 0.76). The fixation count on *Figure* (F_{count}) was statistically significantly different between poor and rich groups $F(1, 69)$ =4.84, $p = 0.03$, $\eta^2 = 0.07$. The mean of total fixation count on the *Figure* was lower in the poor group (*M*= 118, *SD*= 52.7) than the rich group (*M*= 145, *SD*= 52.2), as presented in Table 19. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p =0.03, *d*= -2.20). Therefore, participants' fixation count on the *Figure* was higher when their descriptions vary more.

Table 19
Fixation Counts on Figure

Decription Richness	Mean	SD	N
Poor	118	52.7	37
Rich	146	52.2	34

One-way ANOVA was applied to determine if the fixation counts on the Ground (G_{count}) were different for groups with poor or rich levels of descriptions. There was one outlier in the poor condition, stimuli 24, as assessed by boxplots. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.61). The fixation count on the *Ground* (G_{count}) was statistically significantly different between poor and rich conditions, $F(1, 69) = 12.3$, $p < .001$, $\eta^2 = 0.15$. Further, the mean fixation count on the *Ground* was lower on the poor group (*M*= 83.5, *SD*=46.6) than on the rich group (*M*= 123, *SD*= 48.9), as presented in Table 20. Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant ($p < .001$, d= -3.51). Therefore, participants' fixation count on the *Ground* was more when their description variety was more.

Table 20
Fixation Count on Ground

Decription Richness	Mean	SD	N
Poor	83.5	46.6	37
Rich	123	48.9	34

In sum, participants fixated more and longer on the *Figure* when the description variety of participants was greater. However, their mean fixation duration did not change between groups. In addition, participants fixated more (mean/total) and longer on the *Ground* when the description variety of participants was greater. In other words, except F_{avg}, all dependent variables (F_{sum}, F_{count}, G_{avg}, G_{sum}, G_{count}) showed sensitivity to description richness levels. The following section summarizes the relationship between the lexical variety of utterances and fixations of participants.

4.3.4.Summary

In this part, we analyzed the lexical variety of participants' utterances in terms of fixation data. We considered the description of participants as a whole and divided it into two: spatial terms, reference objects. Data in each group, spatial terms, reference objects, and description, are divided into two conditions: lexically rich and lexically poor. Eye-tracking data on the *Figure* and the *Ground* in the stimulus is compared for each condition (i.e., lexically poor and lexically rich) to see if there are statistically significant differences, i.e., whether topological relations that cause lexically rich utterance also change the ocular behavior of the participants on the *Figure* and the *Ground*.

Overall, results showed that, in general, as participants produced more lexical in the utterance, their fixation duration gets longer, and their fixation counts get more in all three groups. For example, Table 21 shows the fixation counts on *Ground* for the three groups. For three groups: spatial terms, reference objects, and descriptions, the means of fixation count were higher when participants produced lexically rich expressions.

Table 21: Means of Fixation Count on the Ground for Expression Groups

		G_{count}
Spatial Terms	Lexically Poor	86.5
	Lexically Rich	127
Reference Objects	Lexically Poor	87.2
	Lexically Rich	119
Descriptions	Lexically Poor	83.5
	Lexically Rich	123

Results were divided into three for eye tracker variables. Firstly, mean fixation duration on *Figure* and *Ground* (F_{avg} and G_{avg}), total fixation duration on *Ground* (G_{sum}), and fixation count on *Ground* (G_{count}) were more when participants produced lexically rich spatial terms. Secondly, total fixation duration on *Figure* and *Ground* (F_{sum} and G_{sum}), and fixation count on *Figure* and *Ground* (F_{count} and G_{count}) were more when participants produced lexically rich reference object terms. Lastly, total fixation duration on *Ground* and *Figure* (F_{sum} and G_{sum}), mean fixation duration on *Ground* (G_{avg}), and fixation count on *Figure* and *Ground* (F_{count} and G_{count}) were more when participants produced lexically rich descriptions. These results show that when spatial scenes are expressed in various ways with different spatial expressions, the speakers fixated on the *Figure* and *Ground* more, i.e., they fixated longer on the objects.

In the following section, we analyzed the relationship between participants' gaze and words suffixed with locative case markers (LOC).

4.4. Locative Case Markers and Spatial Terms

In section 4.3, spatial expression variety analyses included adpositions and locative suffixes that are attached to words as a variety of spatial terms. However, locative suffixes represent different kinds of topological relationships and do not exist in every language.

The number of the locative suffixes used in each stimulus, rather than adposition, was the bulk of the participants' descriptions. Out of 2528 spatial descriptions in total, locative suffixes occupied approximately 32% (N=817) of the answers. In addition, the analyses of LOC in participants' utterances revealed that to describe 21% (*N*= 15) of the stimuli (TRPS) (*N*=71), participants did not use LOC. For the remaining 79%, at least one participant used locative case markers (LOC). Our results emphasize the common usage of LOC in Turkish. Nevertheless, there are limited studies in the literature on locative case markers in Turkish and spatial language cognition, and most studies focus on spatial terms as adpositions. For example, the spatial configuration of "apple in a bowl" can be represented in two ways in Turkish. Turkish speakers either use adpositions, exemplified in (26), and locative case markers, exemplified in (27) (Sümer et al., 2012).

(26) *kasenin içinde* "in the bowl"

(27) *kasede* "in/on the bowl"

However, most studies on Turkish spatial semantics in the literature included adpositions as in (26) but did not include locative case markers as in (27). Therefore, to assess the importance of locative case markers, we analyzed locative case makers variety to answer the research question "What is the relationship between the variety of locative case markers in Turkish and oculomotor characteristics, response time, and participant's utterance?".

LOC variety was assessed by calculating the number of different words with LOC that were produced for a stimulus. For example, for the stimulus representing dog in a kennel, participants used LOCs suffixed to three different words as in (28). Therefore, the variety of LOC for this stimulus is three.

(28) *Barakada.* "on/in its booth"

Kulübesinde. "on/in its kennel"

Evinde. "in/on its house"

In terms of participants' eye-tracking data, there was no significance between LOC rich and LOC poor conditions. Nevertheless, the following sections present significant differences in number of LOC between the group with rich descriptions and the group with poor descriptions.

4.4.1. Description Richness

Data were classified into two groups: Rich in descriptions and poor in descriptions. There was one outlier in poor condition, as assessed by a boxplot. The data was not normally distributed for the poor groups assessed by the Shapiro-Wilk test (*ps*<0.001), and there was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p= 0.95). The number of locative suffixes (-de/-da) was statistically significantly different between description variety groups, poor and rich, $F(1, 69) = 7.75$, p =0.007, $\eta^2 = 0.10$. The mean of the number of different words with locative suffixes was lower in poor condition (M= 1.95, SD= 1.90) than in the rich condition (M= 3.18, SD= 1.82). Tukey post hoc analysis revealed that the mean increase from poor to rich was statistically significant (p=0.007, d= -2.78). Therefore, participants preferred using locative suffixes in a more various way for the stimulus, for which participants used different descriptions.

Overall, the result of the analyses showed that participants produced more variety of LOC when their description variety was greater, showing the relationship between the variety of LOC and variety of spatial expression.

Further, we analyzed spatial terms that did not involve locative case markers (i.e., adpositions). The results of adpositions revealed that when we exclude LOC from utterance, some statistical significance was lost. The increase in mean fixation duration on the *Figure* when speakers produced richer spatial terms and the increase in mean fixation duration on the *Ground* when speakers produced richer descriptions were not found for adpositions only. Therefore, we found that analyzing adpositions only without including LOC changed the statistical results.

The following section presents the differences in behavioral and linguistic data between *Core* and *Non-Core* groups.

4.5. Core vs. Non-Core Groups of Spatial Scenes

In their research, Anna Papafragou and Barbara Landau (2020), who created a battery to categorize *Containment* and *Support* subtypes, tested *Core* versus *Non-Core* configurations of *Containment* and *Support.* In the *Containment* category, *Full & Loose*, *Full & Tight, Partial & Loose,* and *Partial & Tight* subtypes, and in the *Support* category, *Gravitational Support* subtype were considered *Core*. The rest were considered *Non-Core*. These subtypes and configurations are also studied and accepted by Landau (2020) in cross-linguistic studies.

The distinction between *Core* and *Non-Core* groups was tested in cross-linguistic studies, and it was found that basic spatial expressions, such as *BE on/in* in English and *INE mesa/mesa se* in Greek were used for *Core* groups. Lexically rich expressions such as *hanged on* were used for *Non-Core* groups. The reason for the difference between *Core* and *Non-Core* groups results from function-based prepositions' representation of force-dynamic and mechanical relationships between the *Figure* and the *Ground* in terms of

Topological geometry (Landau, 2020). Function-based prepositions represent properties of openness/closeness and interior/exterior rather than the properties of Euclidian geometry, such as angles and distance. Therefore, other factors, except geometrical properties of the objects, might be involved in the semantics of the preposition. For example, the preposition *on* geometrically represents support from the surface, such as "cup on the table"; however, the expression "spider on the ceiling" holds a force-dynamic relationship between the spider and the ceiling. There is no support from the ceiling to the spider. The relationship is functional, not geometrical. Figure 13 represents two stimuli from our study.

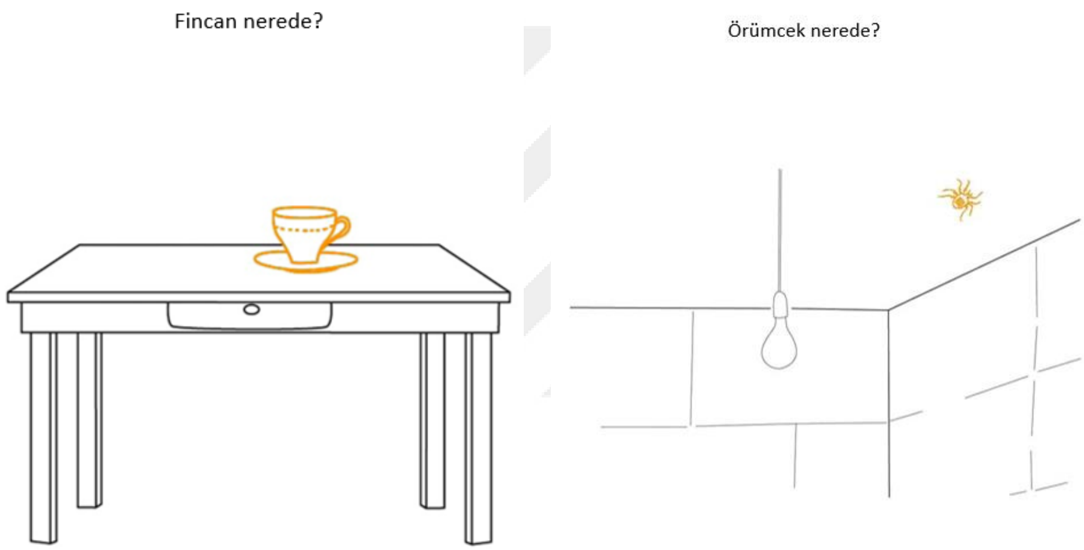

Figure 13: Core and Non-Core examples from the stimuli

Stimuli one, on the right, represents "cup on the table", and stimuli seven, on the right, represent "spider on the ceiling". Stimuli one represents the geometrical relationship of support from below between the table and cup. It was considered as *Gravitational support* subtype of the *Support* type, so it was considered to be in the *Core* group, according to Landau (2017). However, stimulus seven represents the functional relationship between the ceiling and the spider. It was considered as the *Hanging* subtype of the *Support* type. Therefore, stimulus seven was considered to be in the *Non-Core* group, according to Landau. Labeling process of our stimuli as *Core* and *Non-Core* groups was explained later in detail.

Participants used prepositions as in (29) to express the *Core* relationship between the *Figure* and the *Ground* in stimulus one. To express the *Non-Core* relationship between *Figure* and *Ground* in stimulus seven, participants used prepositions as in (30), and also locative case markers as in (31).

(29) *Masanın üzerinde*. "on the table"

Masanın üstünde. "on the table"

Masanın orta sağında. "in middle right of the table"

Masanın sağ üst köşesinde. "in the upper right corner of the table"

(30) *Duvarın üstünde*. "on the wall"

Tavanın üzerinde. "on the wall"

Lambanın yakınında. "close to the lamb"

(31) *Tavanda*. "on/in the ceiling"

Duvarda. "on/in the wall"

Çatıda. "on/in the roof"

Odanın tavanında. "on/in the ceiling of the room"

In the data we have, we labeled the stimulus as *Containment* and *Support types*, divided as *Core* versus *Non-Core* groups. Because there was no literature categorizing the data we have, some assumptions and exclusions were made in the data. The data contains seventy-one stimuli with a variety of relationships between the *Figure* and the *Ground*. 80% (*N*=57) of our stimuli (TRPS) was selected and labeled as *Containment* and *Support* relations. We categorized selected stimuli as *Core* or *Non-Core* according to picture representations examples of Landau et al. (2017).

There were stimuli whose category was unclear. For example, in stimulus 49, it was unclear if it was *Containment* or *Support* because both the expressions 'leaves ON the tree' and 'leaves IN the tree' are used in English. Because in Turkish, 'leaves ON the tree' 'ağaç ÜSTÜNDEKİ yapraklar' is used, Stimulus 49 were taken as *Support* and *Non-Core*. In addition, it was unclear if stimulus 70 can be taken as a *Point Attachment* subtype of *Support* or another relation except for *Support* and *Containment*. It was taken as *Non-Core*; however, when it was excluded from analysis, the p-value for differences between *Non-Core* and *Core* in locative suffixes becomes non-significant. So, the significance should be interpreted carefully.

We analyzed the ocular response of the participants between *Core* and *Non-Core* groups. The hypothesis that there would be significant differences in the ocular response of participants between *Core* and *Non-Core* groups was tested with three MANOVAs and several ANOVAs. MANOVAs were not significant between *Core* and *Non-Core*. Follow-up one-way ANOVA tests were conducted to determine if the mean fixation duration, and fixation count on the *Figure* (F_{avg}, F_{count}), the mean and total fixation duration, and fixation count on the *Ground* (G_{avg}, G_{count}, G_{sum}), and the reference object variety were different for groups between *Core* and *Non-Core*. There was no significance between groups. The data was not normally distributed for F_{avg}, assessed by the Shapiro-Wilk test.

Although there were no significance of other dependent variables between *Core* and *Non-Core* groups, ocular data was significant between *Core* and *Non-Core* groups for total fixation duration in the *Figure* (F_{sum}).

Further, we analyzed the utterance of participants in terms of *Core* and *Non-Core* groups. In terms of linguistic data, spatial terms variety and locative case variety showed significance between *Core* and *Non-Core* groups. In addition, spatial terms variety and description variety displayed significance between *Core* and *Non-Core* groups of *Containment* and *Support*.

The results in the following sections showed there are significant differences between *Core* vs. *Non-Core* groups and *Containment* and *Support* types in terms of behavioral data and lexical analysis. First, participants fixated longer on the *Figure* in total (F_{sum}) for *Non-Core* groups. Second, LOC variety and spatial terms variety increase in *Non-Core* groups. Lastly, description variety was more in *Containment-Non-Core* subtype than *Containment-Core* subtype. These findings imply that participants produce richer spatial expression in *Non-Core* groups. The following sections present detailed analyses and statistical results of the *Core* vs. *Non-Core* group.

4.5.1. Total Fixation Duration on Figure

A one-way ANOVA test was conducted to determine if the total fixation duration on the *Figure* (F_{sum}) differed between *Core* and *Non-Core* groups. Data were classified into two groups: *Core* (24) and *Non-Core* (33). Each group was normally distributed, assessed by the Shapiro-Wilk test. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.42)

Total fixation duration on the *Figure* (F_{sum}) was statistically significantly different between *Core* and *Non-Core* group, $F(1, 55) = 4.71$, p =0.03, $\eta^2 = 0.08$. The fixation duration means was greater on the *Non-Core* group (M= 36583, SD= 13640) than the *Core* group (M=29239, SD= 11188), as presented in Table 22. Tukey post hoc analysis revealed that the mean increase from *Core* to *Non-Core* was statistically significant (p = 0.003, d= -0.58).

Table 22
Total Fixation Duration on Figure

Core vs Non-Core Groups	Mean	SD	N
Core	29239	11188	24
Non-Core	36893	13640	32

4.5.2. Locative Case Marker Variety

One-way ANOVA tests were conducted to determine if the number of locative suffixes attached to different words (LOC) was different for groups between *Core* and *Non-Core*. The data were classified into two groups: *Core* (24) and *Non-Core* (33). Each group was normally distributed, assessed by the Shapiro-Wilk test. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.63)

The number of different words with locative suffixes was statistically significantly different between *Core* and *Non-Core* groups, $F(1, 55) = 4.14$, p =0.05, $\eta^2 = 0.07$. The value of LOC variety was higher in the *Non-Core* group (M= 3.51, SD=1.82) than the *Core* group (M=2.46, SD= 2.08), as listed in Table 23. Tukey post hoc analysis revealed that the mean increase from *Core* to *Non-Core* was statistically significant (p = 0.005, d=-0.546).

Table 23
LOC variety

Core vs Non-Core Groups	Mean	SD	N
Core	2.458	2.085	24
Non-Core	3.500	1.849	32

4.5.3. Spatial Terms Variety

A one-way ANOVA test was conducted to determine if the number of the unique spatial terms differed between *Core* and *Non-Core* groups. Data were classified into two groups: *Core* (24) and *Non-Core* (33). Each group was normally distributed, assessed by the Shapiro-Wilk test. There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.53)

The number of different spatial terms used in stimuli to explain the topological relationship between the *Figure* and the *Ground* was statistically significantly different between *Core* and *Non-Core* groups, $F(1, 55) = 8.39$, p =0.05, $\eta^2 = 0.13$. The number of spatial terms was higher in the *Non-Core* group (M= 8.25, SD=3.26) than the *Core* group (M=5.58, SD= 3.60), as listed in Table 24. Tukey post hoc analysis revealed that the mean increase from *Core* to *Non-Core* was statistically significant (p = .005, d=-2.90).

Table 24
Spatial Terms variety

Core vs Non-Core Groups	Mean	SD	N
Core	5.58	3.60	24
Non-Core	8.25	3.26	32

We further conducted another one-way ANOVA with four groups: *Containment-Core*, *Containment-Non-Core*, *Support-Core*, and *Support-Non-Core* to comprehend the relationship between types. The data were classified into four groups: *Containment-Core* (12), *Containment-Non-Core* (2), *Support-Core* (12), and *Support-Non-Core* (30). There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.25).

The number of different spatial terms used in stimuli to explain the topological relationship between *Figure* and *Ground* was statistically significantly different between *Containment-Core, Containment-Non-Core*, *Support-Core*, and *Support-Non-Core*, $F(3, 52) = 4.77$, $p = 0.005$, $\eta^2 = 0.22$. The number of distinct spatial terms that are used by participants was higher in the *Containment-Non-Core* (*N*=2) group (*M*= 13.5, *SD*=0.71) than the *Containment-Core* (*N*=12) group (*M*=5.50, *SD*= 2.15). Moreover, the number of distinct spatial terms that were used by participants was slightly higher in *Support-Non-Core* (*N*=30) group (*M*= 7.90, *SD*=3.06) than the *Support-Core* (*N*=12) group (*M*=5.68, *SD*= 4.73), as listed in Table 25. Tukey post hoc analysis revealed that the mean increase from *Containment-Core* to *Containment-Non-Core* was statistically significant ($p = 0.01$, $d = -3.17$). However, there was no significance between *Support-Core* and *Support-Non-Core* groups.

Table 25
Spatial Terms variety

Types	Mean	SD	N
Containment-Core	5.50	2.15	12
Containment-Non-Core	13.5	0.71	2
Support-Core	5.68	4.73	12
Support-Non-Core	7.90	3.06	30

4.5.4. Description Variety

Descriptions consist of spatial terms and reference object terms in our analysis. There was no significance between *Core* and *Non-Core* groups in terms of the number of different descriptions participants produced. Therefore, we conducted another one-way ANOVA with four groups: *Containment-Core, Containment-Non-Core, Support-Core,* and

Support-Non-Core to see if there is a relationship between subtypes. Data were classified into four groups: *Containment-Core* (*N*=12), *Containment-Non-Core* (*N*=2), *Support-Core* (*N*=12), and *Support-Non-Core* (*N*=30). There was homogeneity of variances, as assessed by Levene's test of homogeneity of variances (p=0.98). Tukey post hoc analysis revealed that the mean increase from *Containment-Core* to *Containment-Non-Core* was significant (p = 0.05, d= -2.62). The mean description variety for each group is presented in Table 26.

Table 26
Description Variety

Types	Mean	SD	N
Containment-Core	9.92	6.49	12
Containment-Non-Core	21.5	7.78	2
Support-Core	9.42	6.33	12
Support-Non-Core	11.2	5.17	30

4.5.5. Summary

The results showed that there were significant differences between *Core* vs. *Non-Core* groups in terms of behavioral data and linguistic analysis. More detailed analysis with *Containment* and *Support* groups revealed that there were significant differences between *Core* and *Non-Core* subtypes of *Containment*.

Participants' eye movements showed that participants fixated longer on the *Figure* in total (F_{sum}) for *Non-Core* groups. Further, linguistic analyses of spatial terms and locative case markers show that participants tended to use more variety of spatial terms and locative case markers in *Non-Core* groups. When considered the types of *Support* and *Containment*, lexical analysis of the number of different spatial terms and descriptions revealed that participants tended to increase usage of the variety of spatial terms and descriptions from *Containment-Core* subtypes to *Containment-Non-Core* subtypes.

Nonetheless, no differences were found in behavioral data in terms of mean fixation duration on the *Figure* and the *Ground*, fixation count on the *Figure* and the *Ground*, and total fixation duration on the *Ground* (F_{avg}, F_{count}, G_{avg}, G_{count}, G_{sum}) and no differences were found in linguistic analysis in terms of reference objects. The following section presents analyses of the response time as the dependent variable.

4.6. Respond Time

The present thesis' main aim is to use the eye movement of Turkish speakers to study the variety of spatial expressions. The eye-tracking software provides participants' response time in each stimulus, allowing behavioral analysis of response time. However, in the current study, there was no statistical significance in terms of response time.

Several one-way ANOVA tests were conducted to determine whether response time (RT) differed between different lexical richness in spatial terms, locative case markers, descriptions. There was no difference significantly in RT between groups. The same result holds in data without locative case markers.

In addition, a one-way ANOVA test was conducted to determine if response time (RT) was different between *Core* (*N*=24) and *Non-Core* (*N*=33) groups. There was no significance in RT between *Core* and *Non-Core* groups. This result also holds for types of *Containment* and *Support.*

In sum, RT does not change according to the variety of spatial expressions or spatial types, in contrast to the eye movements of participants, which shows the importance of the ocular behavior of participants in spatial scene studies. The following section presents the overall summary of our analyses for the current study.

4.7. Summary of Results

The current thesis' results can be divided into four parts: results on spatial expression variety, locative case markers, *Core* and *Non-Core* types, and response time.

4.7.1. Spatial Expression Variety

We analyzed participants' spatial expression in three groups: spatial terms, reference objects, description.

Our results showed that when participants chose using a variety of spatial terms, in each fixation, participants fixate longer on the *Figure* and the *Ground*. In addition, participants fixated more, spent more time fixating in total on the *Ground* when they chose using a variety of spatial terms. To illustrate, Figure 14 presents two stimuli, stimulus 14 and stimulus 67, from our battery (TRPS).

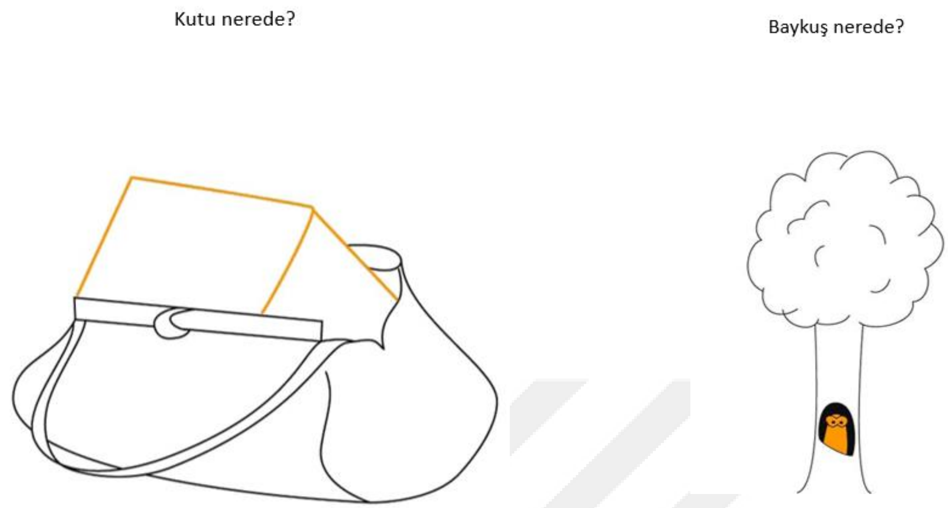

Figure 14: Stimuli in lexically rich and lexically poor conditions

Participants produced various spatial terms, reference objects, and descriptions for the stimulus 67, which presents a spatial relationship between a three *(Ground)* and an owl *(Figure)*. Therefore, stimulus 67 was included *lexically rich* conditions of spatial terms variety, reference objects variety, and descriptions variety.[9] However, stimulus one was included in *lexically poor* conditions of spatial terms variety, reference objects variety, and descriptions variety.

The results showed that participants fixated longer in each fixation on the *Figure* (i.e., the owl) and the *Ground* (i.e., the tree) of stimulus 67 more than the *Figure* (i.e., the box) and *Ground* (i.e., the bag) of stimulus 14. In general, participants fixated longer in each fixation on stimuli which was included in *lexically rich* condition of spatial terms variety than stimuli which was included in *lexically poor* condition of spatial terms variety.

Not only with spatial terms variety but differences of reference object variety also showed differences of eye movements on the *Figure* and the *Ground*. Similarly, participants fixated more and spent more time looking at the *Figure* and the *Ground* for stimulus in which participants used richer vocabulary to express reference objects.

In addition, when the variety of spatial descriptions increased, the duration of fixation in total and fixation count on the *Figure* and the *Ground* increased. Overall, when Turkish

[9] Simuli did not always labeled as *lexically poor* or *rich* for all groups. Stimuli may be included in *lexically rich* condition for spatial terms and in lexically poor conditions for reference objects or descriptions.

speakers produced diverse spatial expressions for a stimulus, their ocular response changed. They fixated more, and their fixation duration got longer. The next section summarizes the results for locative case markers.

4.7.2. LOC variety

Analyses of the linguistic data showed that locatives took a significant place in participants' utterances. Participants preferred using locative suffixes more in the stimulus for which participants used various descriptions; however, there were no differences between locative rich and locative poor groups in terms of the ocular response of participants. The next section summarizes the *Core* vs. *Non-Core* distinction in Turkish.

4.7.3. Core vs. Non-Core Distinction in Turkish

In our study, *Core* and *Non-Core* divisions of spatial expressions were discussed in two types of analyses. First, general distinction of *Core* vs *Non-Core* groups without types was analyzed. In terms of ocular response, Turkish speakers tended to fixate longer on the *Figure* in total. In addition, lexical analyses showed that participants tended to use more variety of spatial terms and locative case markers for *Non-Core* groups.

Second, a more detailed analysis including types, i.e., *Containment-Core*, *Containment-Non-Core*, *Support-Core*, and *Support-Non-Core*, revealed that Turkish speakers produced more lexically rich spatial terms and descriptions when they saw the *Containment-Non-Core* subtype stimuli than *Containment-Core* subtype stimuli.

4.7.4. Response Time

There was no statistical significance of participant’s response time between any group.

The following chapter presents discussions about the results, limitations, and possible future work.

CHAPTER 5

DISCUSSION AND FUTURE WORK

5.1. Discussion

The semantics of spatial language and conceptualization of space has been widely studied in other languages by investigating cross-linguistic similarities to understand variations of spatial expressions between languages. Nevertheless, Johannes et al. (2015) suggested that cross-linguistic similarities can also be investigated by assessing within language variation. Therefore, in this study, we analyzed within language variation of Turkish spatial language semantics.

In the current study, we presented detailed analyses of the semantics of spatial relations in Turkish with language and eye-tracking data. We found significant relationships between language and ocular response along with language and spatial term types. In addition, we emphasized the importance of including locative case markers in Turkish spatial language studies.

Summaries of linguistic and ocular analyses are given in chapter 4. The following sections discuss the implications of lexical variety, locative case markers, spatial terms categorization, and response time results.

5.1.1. Degree of lexical diversity and its ocular implication

In Turkish, speakers use various spatial expressions, including words with locative case markers, to express topological relationships between the *Figure* and the *Ground*. Our study investigated the relationship between the ocular response of Turkish speakers and their spatial expression variety. The results of our study on spatial expression diversity have shown a significant relationship between Turkish speakers' production of spatial expression variety to describe the stimuli (TRPS) and their ocular response when they look at the stimuli. Their ocular response had a significant relationship with the variance of spatial expressions.

Gaze movements are indicators of cognitive processing. Studies have shown that as fixation gets more and longer, observers' visual attention increases (Henderson, 2011). In addition, the visual and cognitive difficulty of a scene results in longer gaze durations

(Henderson et al., 2014). Research on the spatial orientation of objects revealed that eye movements are influenced by the spatial orientation of objects (Cronin & Brockmole, 2016). Research on visual perception and eye movements shows that cognitive processing is strongly related to fixation duration and supported the findings in our study.

Analysis of the lexical and ocular data revealed that, in general, if lexical variety of spatial expressions participants produced was high, participants' fixation count, duration, and mean fixation duration were higher on both *Figure* and *Ground*. Prolonged fixation duration and increase in fixation count imply cognitive demand and visual attention. These results show the effect of spatial scenes that provide rich spatial expressions on cognitive processing, causing prolonged and higher gaze movements.

Roundup results showed the gaze duration and count were higher on the *Figure* and the *Ground* when the stimulus is expressed with rich vocabulary instead of poor vocabulary; however, there were some insignificances between groups of lexical richness levels when looked at in detail. We analyzed participants' utterances in three groups: spatial terms, reference objects, descriptions. In the analyses of spatial term variety, the level of lexical richness of spatial terms did not affect the total fixation duration and counts on the *Figure* (F_{sum} & F_{count}). In contrast, it affected the mean fixation duration on the *Figure* (F_{avg}). In addition, we got the exact opposite results in reference object variety; the level of lexical richness of reference object did not affect mean fixation duration on the *Figure* (F_{avg}) and the *Ground* (G_{avg}), while it affected the total fixation duration and fixation count on the *Figure* and the *Ground* (F_{sum}, F_{count}, G_{sum}, and G_{count}). Similarly, the level of lexical richness of descriptions did not affect mean fixation duration on the *Figure* (F_{avg}), while it affected the total fixation duration and fixation count on the *Figure* (F_{sum} and F_{count}). Hooge et al. (2007) reported that mean fixation duration is found in more difficult tasks; however, they also noted that this is not a one-to-one relationship. Therefore, there were no clear findings of cognitive reasons of differences between mean fixation duration and fixation count of observers. Yet, it was seen in our study that an increase in fixation count and mean fixation duration does not come in together.

The average number of descriptions did not change significantly between stimuli conditions, so the reason for more fixation count and longer duration on some stimuli were not caused by the number of the descriptions that participants produced; rather, it was the lexical variety that participants produced. As a result, the topological relations with conceptual complexity result in spatial expression variety, inducing longer lexical retrieval.

5.1.2. Importance of Locative Case Markers

In Turkish, spatial relationship is expressed with locative case markers (LOC) along with spatial terms. In the literature, spatial cognition was studied with LOC rarely, and it is found that Turkish speakers preferred LOC rarely (Sümer et al., 2012). Nevertheless, locative case markers occupied approximately 32% of our data. In addition, for 79% of the stimuli (Topological Relations in Picture Series, Bowerman & Pederson, 1992), at least one participant used locative case markers. For this reason, it is essential to include locative case markers in the studies, yet the research on this topic is weak. The importance of locative case markers was investigated by analyzing locative case markers and adpositions separately. We found that if the value of LOC variety is high for a stimulus, the value of description variety is also high. This result shows that participants used more various LOC when they used more various descriptions.

Further, the analysis of adpositions exposed that without locative case markers in the data, some eye tracking results disappeared. Considering the effect of spatial terms and description variety on gaze duration is one of the main findings of the current study, we believe that locative case markers influence results and inferences of studies cognition of spatial expressions.

Moreover, analysis of *Core* and *Non-Core* categorization revealed that participants used LOC variety more on *Non-Core* groups. These results show the importance of locative case markers also in *Core* categorization studies.

Overall, studies should include not only adpositions but also markers in Turkish spatial term studies. Including spatial markers may also change the results of cross-linguistic studies. The next section discusses *Core* and *Non-Core* distinctions in Turkish.

5.1.3. Within-Language Variety and Spatial Categorization

Landau (2017) set types of *Support* and *Containment* apart from other spatial types such as *Occlusion* because of their force-dynamic nature. Landau claims that the reason for cross-linguistic variety in *Support* and *Containment* can be understood if the cultural and functional inference on using *Support* and *Containment* adpositions is considered. In addition, the types of *Support* and *Containment* are also divided into *Core* and *Non-Core* groups, including types of spatial configurations. *Core* groups include geometrical spatial configurations, such as *Gravitational Support* subtype of *Support* type and *Full Containment* subtype of *Containment* type. Whereas¸ *Non-Core* groups include function-based spatial configurations such as *Hanging* subtype of Support and the *Embedded* subtype of *Containment* type.

Research have been conducted to understand the use of basic expressions *BE in/on* in *Core* and *Non-Core* groups and if similarities hold cross-linguistically (Johannes et al., 2015; Landau, 2017, 2020; Landau et al. 2017). Studies revealed that although there are differences between languages, basic expressions are found in *Core* groups more than *Non-Core* groups, and the variety of expressions increase in *Non-Core* groups. However, not all studies support the idea of dividing all languages as *Core* and Non-*Core* prior to the study. Feist and Zhang (2019) reported that some studies classified the scenes as exemplars of *Containment* and Support (Johannes et al. 2015; Landau et al. 2017) despite the cross-linguistic evidence on the contradiction of coherence of *Support* (Levinson & Meira, 2003). It has been shown that *Support* relations are also clustering with two different groups of scenes. Feist and Zhang (2019) concluded in their study that cultural complexity is added to universal *Core* concepts.

In the current study, we used the stimuli of Landau et al. (2017) to categorize our battery into types and *Core* vs. *Non-Core* groups. Then we analyzed ocular and linguistic data, divided into types to test claims of studies supporting a priori type division from a different point of view. The analyses of participant's utterances revealed that Turkish speakers used more various spatial terms to describe the function-based spatial configuration in *Non-Core* groups than *Core* groups. Similarly, Turkish speakers produced lexically rich descriptions and lexically rich spatial terms for the *Non-Core* group of *Containment* type stimuli more than the the *Core* group of *Containment* type stimuli. Moreover, one of the important result of our study is the LOC variety. Our study showed that Turkish speakers used more various LOC for *Non-Core* groups. The reason for that might be the nature of locative case markers. The locative case markers (LOCs) do not reveal the exact nature of relationships between the *Figure* and the *Ground*.

These findings support the claims of Landau (2020) on the generalizability of *Core* vs. *Non-Core* categorization across languages from a different point of view. In the related studies, each stimulus with spatial configurations was labeled most used spatial terms, and the between-language variety of expression and similarity between languages were assessed. However, in the current study, we calculated the variety of spatial expressions produced by participants for each stimulus and assessed if there is a difference between *Core* and *Non-Core* groups in terms of expression richness. Our results supported the *Core* vs. *Non-Core* distinction in Turkish.

Ocular data revealed that the differences between *Core* and *Non-Core* groups affect total fixation duration on the *Figure*, yet, mean fixation duration and fixation count have not been affected by the *Core* vs. *Non-Core* categorizations. In addition, fixation duration and fixation counts on the *Ground* were not affected by *Core* categorization at all. This result is different from the results of the general lexical richness analysis. Fixation data on the *Ground* was affected as lexical richness increase in spatial terms, reference objects, and descriptions; however, fixation count and duration on the *Ground* did not change between *Core* and *Non-Core* groups even though spatial terms and locative case variety increased. The fact that only fixation duration on *Figure* showed sensitivity to *Core* vs. *Non-Core* categorization might give information about the distinction of *Support* and *Containment* types from other groups at a cognitive level. A more detailed study should be conducted to understand this difference.

5.1.4. Importance of Ocular Data

Studies on non-linguistic tests showed that linguistic data alone could not provide enough information about patterns of spatial cognition (Le Guen, 2011). It is essential to use non-linguistic methods to understand the semantic categorization of spatial terms. In this study, along with linguistic data, the ocular response of participants provided information on spatial cognition. Using linguistic and non-linguistic data together helps us to understand spatial cognition better. Fixation data revealed the cognitive demand for lexically rich spatial expressions and *Non-Core* groups, as well as the importance of locative case markers. In addition to the ocular response, eye trackers provide participants' response times, a vital variable in psychology studies. However, neither spatial expression variety nor *Core* vs. *Non-Core* categorization affects the response time of participants, which reveals that if we did not gather ocular information of participants, we would not infer differences with response time only. More studies should be done using different behavioral assessments such as eye movement to comprehend spatial cognition in detail.

5.2. Limitations and Future Studies

In this thesis, we have analyzed Turkish spatial expression data and speakers' ocular data. We considered locative case markers and subtypes of spatial scene types of Landau (2017, 2020) categorized. Comparison of analyses in terms of subtypes of spatial types gave indefinite information about the effects of locative case markers on each subtype. The battery is not balanced in terms of subtypes of spatial types. For example, while the battery has 12 *Containment-Core* type stimuli, there are only two *Containment-Non-Core* stimuli. These unbalanced numbers affect the liability of statistical analysis. In addition, we conducted an open-ended study for linguistic freedom. However, a fixed task, in which the participants are restricted to use terms from a limited set for locative case markers, may give more detailed results. More studies should be done to understand the subtypes of spatial types in Turkish with a more suitable battery. Moreover, there are not many studies on cognition of Turkish spatial language in the literature. More studies, including locative case markers, should be conducted in the future.

Cross-linguistic studies, including Turkish, are limited, and the ones including Turkish lacks enough participants and spatial expression variety, as mentioned before. Johannes et al. (2015) conducted a cross-linguistic study to show the universality of variation in the number of expressions for different types of relations. They worked on a broad category of spatial scenes rather than in each stimulus. They revealed that the variation in spatial expression in each subtype is similar across languages, different than our study. Future studies on Turkish spatial language should include different languages and analyze each stimulus to understand the universality of spatial expression variety.

In the study of Landau et al. (2014), speakers of English answer the question "Where is the figure?" in an inconvenient way, such as "the sock tore" instead of "the hole is in the sock.". Similar expressions can be seen in Turkish spatial expressions too. In the current study, there were some expressions, such as *elmanın içinden geçmiş* 'gone through the apple' instead of *elmanın içinde* 'in/inside the apple'. However, Greek speakers did not produce such utterances, so Landau et al. (2014) explained the differences between English and Greek as English being satellite frame and Greek being verb-framed, as mentioned before. If this was the case, we should not have seen these expressions in

Turkish because Turkish is also a verb-farmed language. Therefore, future studies should investigate more to find the reason for the differences.

Landau (2017) claims that the reason for differences in spatial terms between languages is the nature of force dynamic relationships and their limitless functional properties that are not related to geometrical properties. In the current thesis, the results have shown that the variety of the spatial expressions is more in *Non-Core* groups than *Core* groups. We also see ocular data changed as the variety of spatial expression gets more. Here we asked a research question about a variety of spatial expressions within language rather than between languages. Can we relate the increase in variety of spatial expressions in *Non-Core* groups to force dynamic and functional nature of spatial expressions of *Containment* and *Support*? Future studies should investigate the relationship between force dynamic relationships and within language variety.

The main purpose of the current study is to reveal the relationship between eye fixations and spatial expressions in Turkish. It is found that stimulus, expressed in rich vocabulary, affects the fixation duration and count on the *Figure* and the *Ground.* Still, the change in gaze movements is not clear in the current study because of the nature of the setup. Future studies should focus on the nature of change in the gaze movements to reveal the specific reason for the difference.

REFERENCES

Arık, E. (2017) Türkçede Yer, Yönelim ve Hareket Tarzının Kodlanması. *Ulusal Dilbilim Kurultayı Bildirileri*, 61

Atak, A. (2018). Türkçede Uzamsal Dilin Konumlanış Açısından İncelenmesi (PhD thesis, Ankara Üniversitesi, Ankara, Türkiye) Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp,

Barthelmé, S., Trukenbrod, H., Engbert, R., & Wichmann, F. (2013). Modeling fixation locations using spatial point processes. *Journal of Vision*, *13*(12), 1–34. https://doi.org/10.1167/13.12.1

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review*, 94(2), 115.

Bowerman, M. (1996). Learning how to structure space for language: A crosslinguistic perspective. In P. Bloom, M. A. Peterson, L. Nadel, & M. F. Garrett (Eds.), *Language and space,* 385–436. Cambridge, MA: MIT Press

Bowerman, M., & Pederson, E. (1992). Topological relations picture series. In Space stimuli kit 1.2 (p. 51). Max Planck Institute for Psycholinguistics.

Carstensen, A., Kachergis, G., Hermalin, N., & Regier, T. (2019). “Natural concepts" revisited in the spatial-topological domain : Universal tendencies in focal spatial relations. *Proceedings of Cognitive Science* 2019, 197–203.

Carlson, L. A., & Hill, P. L. (2007). Experimental methods for studying language and space. Methods in cognitive linguistics. Amsterdam: John Benjamins Publishing Company. 250-276.

Cronin, D. A., & Brockmole, J. R. (2016). Evaluating the influence of a fixated object's spatio-temporal properties on gaze control. *Attention, Perception, and Psychophysics*, *78*(4), 996–1003. https://doi.org/10.3758/s13414-016-1072-0

Dasen, P. R. (2018). Cross-cultural research on spatial concept development. Cognitive processing, 19(1), 93-99.

Duchowski, A. T. (2017). Eye Tracking Methodology: Theory and Practice. In *Eye Tracking Methodology*. https://doi.org/10.1007/978-3-319-57883-5

Evans, K., Rotello, C. M., Li, X., & Rayner, K. (2009). Scene perception and memory revealed by eye movements and receiver-operating characteristic analyses: Does a cultural difference truly exist? *The Quarterly Journal of Experimental Psychology*, *62*(2), 276–285. https://doi.org/10.1080/17470210802373720

Feist, M. I., & Zhang, Y. (2019). Mapping Space: A Comparative Study. In *CogSci* (pp. 1717-1723

Furman, R., Küntay, A. C., & Özyürek, A. (2014). Early language-specificity of children's event encoding in speech and gesture: evidence from caused motion in Turkish. *Language, Cognition and Neuroscience*, 29(5), 620-634.

Furman, R., Ozyurek, A., & Küntay, A. C. (2010). Early language-specificity in Turkish children's caused motion event expressions in speech and gesture. In 34th Boston University Conference on Language Development [BUCLD 34], 126-137. Cascadilla Press.

Göksel, A., & Kerslake, C. (2005). Turkish: A comprehensive grammar. Routledge.

Henderson, J. M., Olejarczyk, J., Luke, S. G., & Schmidt, J. (2014). Eye movement control during scene viewing: Immediate degradation and enhancement effects of spatial frequency filtering. *Visual Cognition*, *22*(3), 486–502. https://doi.org/10.1080/13506285.2014.897662

Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin and Review*, 15, 566573

Henderson, J. M. (2011). Eye movements and scene perception. In Liversedge, S., Gilchrist, I., & Everling, S. (Eds.). The Oxford handbook of eye movements. (pp. 593-602) Oxford University Press.

Henderson, J.M. and Ferreira, F. (2004) Scene perception for psycholinguists. In The Interface of Language, Vision, and Action: Eye Movements and the Visual World (Henderson, J.M. and Ferreira, F., eds), Psychology Press

Henderson, J. M., Weeks Jr, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology*: Human perception and performance, 25(1), 210.

Henderson, J. M., & Hollingworth, A. (1998). Eye movements during scene viewing: An overview. Eye guidance in reading and scene perception, 269-293.

Henderson, J. M., & Smith, T. J. (2009). How are eye fixation durations controlled during scene viewing? Further evidence from a scene onset delay paradigm. *Visual Cognition*, 17(6-7), 1055-1082.

Henderson, J. M. (2017). Gaze control as prediction. *Trends in Cognitive Sciences*, 21(1), 15-23.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 7(11), 498-504.

Heine, B. (1997). Cognitive foundations of grammar. New York: Oxford University Press.

Hooge, I. T. C., Vlaskamp, B. N., & Over, E. A. (2007). Saccadic search: On the duration of a fixation. In *Eye Movements* (pp. 581-595). Elsevier.

Johannes, K., Wang, J., Papafragou, A., & Landau, B. (2015). Similarity and Variation in the Distribution of Spatial Expressions Across Three Languages. In *CogSci.*

Johanson, M., & Papafragou, A. (2014). What does children's spatial language reveal about spatial concepts? Evidence from the use of containment expressions. *Cognitive Science*, 38(5), 881-910.

Johnston, J. & Slobin, D. (1979). The development of locative expressions in English, Italian, Serbo-Croation, and Turkish. *Journal of Child Language*, 6, 529-45.

Lakusta, L., Brucato, M., & Landau, B. (2020). Evidence for a Core Representation for Support in Early Language Development. *Language Learning and Development*, 16(2), 180-195.

Landau, B. (2017). Update on “what” and “where” in spatial language: A new division of labor for spatial terms. *Cognitive Science*, 41, 321-350.

Landau, B. (2020). Learning Simple Spatial Terms: Core and More. *Topics in Cognitive Science*, 12(1), 91–114. https://doi.org/10.1111/tops.12394

Landau, B., & Jackendoff, R. (1993). " What" and" where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, 16, 217-217.

Landau, B., Johannes, K., Skordos, D., & Papafragou, A. (2017). Containment and Support: Core and Complexity in Spatial Language Learning. *Cognitive Science*, 41, 748–779. https://doi.org/10.1111/cogs.12389

Le Guen, O. (2011). Speech and gesture in spatial language and cognition among the Yucatec Mayas. *Cognitive Science*, 35(5), 905-938.

Levinson, S.C. (2003). Space in Language and Cognition. Cambridge: Cambridge University Press.

Levinson, S. C., Meira, S., & The Language and Cognition Group (2003). "Natural concepts" in the spatial topological domain–adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language*, 79(3), 485–516.

Li, P., & Abarbanell, L. (2018). Competing perspectives on frames of reference in language and thought. *Cognition*, 170, 9-24.

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.

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: a computational model of fixation durations in scene viewing. *Psychological Review*, 117(2), 382.

Rayner, K., Castelhano, M. S., & Yang, J. (2009). Eye movements when looking at unusual/weird scenes: Are there cultural differences?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 254.

Sümer, B., Perniss, P. M., Zwitserlood, I. E. P., & Özyürek, A. (2014). Learning to express" left- right" &" front-behind" in a sign versus spoken language.

Sümer, B., Zwitserlood, I. E. P., Perniss, P. M., & Özyürek, A. (2012). Development of locative expressions by Turkish deaf and hearing children: are there modality effects?.

Talmy, L. (2000a). Toward a Cognitive Semantics, vol. 1: Concept Structuring Systems. Cambridge MA: MIT Press.

Toplu, A. B. (2011). Linguistic expression and conceptual representation of motion events in Turkish, English and French: An experimental study. Unpublished PhD Dissertation). Middle East Technical University, Ankara

Võ, M. L. H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision*, 9(3), 24-24.

Yun, H., & Choi, S. (2018). Spatial Semantics, Cognition, and Their Interaction: A Comparative Study of Spatial Categorization in English and Korean. *Cognitive Science*, 42(6), 1736–1776. https://doi.org/10.1111/cogs.12622

Žilinskaitė-Šinkūnienė, E., Šķilters, J., Zariņa, L., & Bērziņa, N. (2019). Containment and support: similarities and variation in Lithuanian, Latvian and Estonian. *Baltistica*, 54(2), 205-255.

Žilinskaitė-Šinkūnienė, Eglė, Jurģis Šķilters, Līga Zariņa (2019). Containment and support in Baltic languages: Overview, experimental evidence, and an extended RCC as applied to Latvian and Lithuanian, *Baltic Journal of Modern Computing* 7(2), 224–254 . https://doi.org/10.22364/bjmc.2019.7.2.02

APPENDICES

APPENDIX A

The Instruction Page

Deneyin bu bölümünde üzerinde bir soru olan resimler göreceksiniz, aşağıdaki gibi.

Her resimde bir sarı obje var (bu örnekteki kedi), üstteki soru her zaman bu objenin nerede olduğunu soruyor. Sizden istediğimiz soruyu sesli olarak cevaplamanız (mesela «kedi masanın altında»). Bunun gibi tam ve anlaşılır cümleler kurmanız önemli. Toplam 70 civarında resim var. Çalışmanın hızlı olması için söylediklerinizi kaydedip sonra yazıya dökeceğiz. Bu nedenle mikrofon takmayı unutmayın.

Şimdi bir tuşa bastığınızda ilk resim gelecek. Cevabınızı söyleyin ve sonraki resme geçmek için yine klavyeden bir tuşa basın.

APPENDIX B

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

Sayı /

Konu : Değerlendirme Sonucu

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi : İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Doç.Dr. Cengiz ACARTÜRK

Danışmanlığını yürüttüğünüz yüksek lisans öğrencisi Şeyma Nur ERTEKİN'in "Türkçe'de Static Uzumsal Dil Semantiğinin Kelime Çeşitliliği Açısından İncelenmesi: Bir Göz Hareketleri Takip Çalışması" başlıklı araştırmasının ön değerlendirmesi sonucunda araştırmanın arşiv araştırması olduğu görülmüş olup İnsan Araştırmaları Etik Kuruluna başvurmasına gerek olmadan araştırmanın yapılması uygun bulunmuştur.

Saygılarımızla bilgilerinize sunarız.