LOOKING BIAS: AN EXAMINATION OF THE RELATIONSHIP BETWEEN VISUAL SEARCH AND PRENOMINAL ADJECTIVE ORDER IN ENGLISH

A Thesis
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
MASTER OF ARTS

---

by
Victoria R. Rodriguez
May 2017

Thesis Approvals:

Dr. Jamie Reilly, Thesis Advisor, Department of Communication Sciences and Disorders
Dr. Jinyi Hung, Department of Communication Sciences and Disorders

ABSTRACT

The aim of the present study was to determine the relationship between the order of analysis of objects within the visual system and prenominal adjective ordering rules in English, as past syntactic and semantic theories have proven insufficient to explain the phenomenon in its entirety. Three experiments were designed to investigate whether ordering preferences when multiple adjectives are stacked before a noun are determined by properties of the visual system that subsequently map directly onto language via the semantic system. First, an experimental protocol was designed to discover whether participants' visual search pattern varied based on the type of stimuli presented. A second experiment was created to determine whether participants observed features of objects in an order that corresponded to grammatical adjective ordering rules in English. A third and final experiment was devised to explore whether inversions of adjective categories typically positioned closer to the noun were more acceptable than inversions of adjective categories placed further away from the noun or vice versa. Eye tracking data was analyzed for scan sequence (Experiments 1 and 2) and acceptability judgments were obtained using a 7-point Likert Scale survey (Experiment 3). Results showed that participants did not vary systematic scan patterns based on image type, with a greater propensity to not fixate when presented with shapes. Data from the second experiment demonstrated that participants viewed objects in an order that was correlated with prenominal adjective ordering with varying levels of significance. Acceptability judgments from the third experiment indicated that inversions of adjective classes that are typically placed closer to the noun were generally more acceptable than inversions of adjective classes typically placed further from the noun. This study provides preliminary evidence that language rules may be derived from properties of the visual system and cognition. Further research is necessary to explore the nature and extent of correlations between perception, the semantic system, and grammatical features of language.

DEDICATION

To Nikhil Philip and Linda Young.

Thank you both, for everything.

ACKNOWLEDGEMENTS

This text was completed thanks to the support and encouragement of a number of individuals, and to give them all adequate thanks within a single page is an impossible task. Nevertheless, I shall make my best attempt. First and foremost, I would like to thank Amanda Eiser, for bringing each of our “blobs” into existence, for running participants, and for her assistance in data processing. Her persistent and vigorous approach to the development of the experimental protocols were crucial throughout the early phases of this investigation. I am beyond grateful for her input into this project. I must also extend my appreciation to the other members of the Eleanor Saffran Center for Cognitive Neuroscience, many of whom have become friends over the three years that I spent as a volunteer. In particular, I would like to acknowledge Victoria Diedrichs, Bonnie Zuckerman, Jameer Hart, Richard Binney, Helen Felker, Allie Kelly, and Hillary Waller. Through extended discussions and thoughtful advice, you have made great contributions to each phase of this project. Finally, I am particularly grateful for the guidance of Dr. Jamie Reilly and Dr. Jinyi Hung, of the Memory, Concepts, and Cognition Laboratory. This effort was made possible with your help. Thank you for your collaboration in the development of this study, and for your input throughout its implementation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Cross-linguistically, it has been observed that the order of multiple adjectives stacked before the noun (i.e., prenominal adjective order) is subject to certain rules that determine which modifiers are positioned closest to the noun. Previous literature attempted to explain why certain classes of adjectives appear closer to the noun, and why some are positioned further away. However, current syntactic, semantic, and cognitive theories are not sufficient to determine a precise ordering of multiple adjectives that occur prior to the noun in English. An overview of the rules for prenominal adjective ordering in English will be presented, along with descriptions of semantic and syntactic theories that aim to explain adjective sequencing. Following a review of the relationship between the visual and semantic systems, and the implications that eye-tracking research may have for language processing, the potential for feature salience within the visual system as an explanation for the phenomenon of prenominal adjective sequencing will be discussed.

Overview of Prenominal Adjective Order in English

In the fall of 2016, BBC editor Matthew Anderson tweeted the following excerpt from *The Elements of Eloquence,* a 2013 text on proper English rhetoric by English etymologist Mark Forsyth:

> Adjectives in English absolutely have to be in this order: opinion-size-age-shape-colour-origin-material-purpose Noun. So you can have a *lovely little old rectangular green French silver whittling knife*. But if you mess with that word order in the slightest you'll sound like a maniac. (pp. 45-46)

The message was discussed by Ari Shapiro and Audie Cornish of National Public Radio ("BBC Language Rule," 2016), and it has been retweeted an additional 53,000 times since it was first written. Forsyth and other grammarians would argue that these adjective ordering rules, which are intuitive for native speakers, are irrefutable, with different adjectives all belonging to set categories (or classes) that are organized in a hierarchical sequence prior to a noun. However, analysis of the current sources created to aid English language learners in acquiring rules for prenominal adjective ordering would reveal that that these rules are not always clear or consistent, as exhibited by the following comparison in Table 1.

Table 1. A Comparison of Adjective Ordering Conventions in English

Adjective Position	UVELC	Cambridge Dictionary	British Council	Education First
Adjectives farthest from the NP	Opinion	Opinion	General Opinion	Quantity
	Size	Physical Quality	Specific Opinion	Value/Opinion
	Age	Shape	Size	Size
	Shape	Age	Shape	Temperature
	Color	Color	Age	Age
	Origin	Origin	Color	Shape
	Material	Material	Nationality	Color
	Purpose	Type	Material	Origin
Adjectives closest to the NP		Purpose		Material

Note. Opinion adjectives include value judgments (e.g., good, bad); Origin and nationality categories describe place of birth (e.g., American); Material adjectives describe composition (e.g., wool). It should be noted that not all the websites cited above provide example words or that correspond to each adjective class presented.

For example, The University of Victoria English Language Center (“UVELC,” 2011) dictates the following sequence before the noun phrase (NP): opinion > size > age > shape > color > origin > material > purpose. The sequence is almost the same as that prescribed by the Cambridge Dictionary (n.d.), except for the addition of the adjective classes of “type” and “physical quality” (without providing examples that fit membership of these categories). The website for the Cambridge Dictionary further states that adjectives that describe opinions or attitudes are usually located a greater distance from the noun, while adjectives that are factual tend to land closer to the noun. The British Council (n.d.), an international organization that promotes education in English, created a general/specific distinction within the otherwise universal adjective class of “opinion,” while the website of Education First, which provides academic training for multinational citizens, included a separate and unique distinction for an adjective class containing modifiers that describe “temperature” (2017).

While there are several similarities between prenominal adjective ordering rules (e.g., adjectives of opinion are further from the noun than nouns that denote shape or color), the above examples demonstrate that there are obvious differences regarding which adjective classes take precedence when stacking two or more adjectives before a nominal phrase. Additionally, there is disagreement as to how many adjective classes are necessary for the categorization of all nouns that exist in the English language. According to Wullf (2003), there is substantial disagreement as to the number of classes that would be required to classify all English adjectives. These prenominal rules designed to instruct non-native speakers of English do not address abstract adjectives that may not meet any of the prescribed criteria (ex., *punctual* or *geographical*). Further, there are specific situations where it is appropriate to switch an expected, or canonical order, such as instances where stress is used by speakers to make a distinction between two objects.

For example, the phrase **The brown small dog* could theoretically be used to distinguish between multiple objects bearing the attribute of “cuteness,” even though most sources would place adjectives of opinion/value prior to adjectives of color. In a 1972 study, Danks and Schwenk found that inversions of canonical adjective positioning rules could occur in the case of size and color adjectives stacked prior to a noun (where size adjectives typically appear further away than color adjectives) could be judged as more acceptable if they appeared in contexts where color was a distinguishing attribute. In other situations, an inversion of canonical order (i.e., a non-canonical pair) is absolutely unacceptable. An example is apparent in the BBC article cited at the beginning of this paper, where **My Greek, Fat, Big Wedding* cannot be used as an alternative description for *My Big, Fat, Greek Wedding* (“BBC Language Rule,” 2016). Even if there are rules to be derived for prenominal adjective sequencing, these rules can shift based on a variety of factors.

A series of different theoretical models have been proposed to explain the rules behind prenominal adjective sequencing cross-linguistically. Early research concerning prenominal adjective arrangement in English hoped to explain the phenomena in terms of syntactic regulation (Annear, 1965; Vendler, 1967). Later studies provided evidence for a semantic theory of prenominal adjective ordering. What follows is a review of grammatical and semantic theories of prenominal adjective order, as well as a possible cognitive-semantic explanation which formed the basis for this paper.

Syntactic Theories of Adjective Sequencing

The earliest studies of prenominal adjective order relied on a syntactic explanation. Most of these syntactic accounts relied on the idea of different adjective "classes," whose order is determined by an underlying grammar. The idea of the existence of universal forms underlying grammatical structures across all languages, or universal grammar, was revived in modern linguistics by Noam Chomsky in the 1950s (see Chomsky, 1959; Chomsky, 1956). At this point in time, researchers and linguists alike attempted to discover evidence that a generative grammar, and transformations to that generative grammar, could lead to uniform superficial forms cross-linguistically. Below are accounts from several researchers that attempted to use transformations to derive rules for prenominal adjective order.

In a 1964 paper, Sandra Annear tried to predict adjective order by giving 3 order classes of adjectives. This order was dictated by determining whether the modifiers examined belonged to class 1, class 2, or class 3. Classes were defined by morphological markers; for example, one class consisted of modifiers that could have the adverbial marker –LY added at the end. Her research presupposed that adjective order rules could be derived by transformations in a three-stage route, consisting of one embedding transformation and two single-based transformations (Annear, 1964). An example from Annear's three-stage route is outlined below:

> Det + N + Tns + BE + ADJ is the general terminal string:
> *The boy is tall*
> A WH* marker is attached to boy and embedded into the string:
> *The boy is my brother* → *The boy who is tall is my brother*
> The relative clause reduction rule gives:
> *The boy tall is my brother*
> And application of the transposition rule gives:
> *The tall boy is my brother.* (pp. 96-97)

Although Annear attempted to outline a route by which multiple adjectives occurring before the noun could be organized, she stated that there was no other function in language for these classes except to sequence multiple adjectives when they occur before the noun (1964). Thus, she concluded that adjectives derived from syntactic transformations was an unlikely explanation for adjective sequencing, and that adjective order is likely determined by factors outside of the realm of syntax. In his paper, Sussex similarly proposed that the use of classes to determine adjective sequence was insufficient because they created rules that were not applicable to any other area of language (1974). Ultimately, a generative view of prenominal adjective ordering has largely fallen out of favor with linguists, along with theories of generative grammar. In its place, semantic theories arose to explain the phenomenon.

Semantic Theories of Adjective Sequencing

It was found that early theories, which relied on separating adjectives into classes that had a set order prior to the noun, were insufficient to explain the rules of prenominal adjective order in English. In a 1969 paper, Martin suggested that the order classes proposed by earlier experiments were indistinguishable from semantic classes. Based on data collected in a series of experiments, Martin introduced semantic closeness, or definiteness, as the determiner of prenominal adjective order as a part of a semantic theory of adjective sequencing. Specifically, adjectives more similar to the noun in meaning would be closest to the noun in position. As a part of his research, Martin presented evidence that definiteness of denotation was correlated most significantly with adjective order. In his account, adjectives that describe value or size are dependent on comparative relationships between words, while adjectives that denote age or color retain their meaning regardless of the noun phrase that they modify. Thus, the adjective *big* is less definite than the adjective *gray* because *big* can describe a property that varies significantly across objects (a big skyscraper is significantly larger than a big dog), while the meaning of the modifier *gray* is stable regardless of the object it modifies. It is these most definite features that are positioned closer to the noun when multiple descriptors are combined. This finding was supported by evidence from Danks and Glucksberg (1971), which showed that adjectives that were placed closer to the noun in position were likely to indicate a more intrinsic, innate property about the noun.

A study conducted by Scontras, Degen, and Goodman (2017) argued that preferred adjective sequences in English and other languages with both prenominal and post-nominal ordering rules may be governed by general properties of cognition, namely a “universal hierarchy” of semantic classes that would determine order in situations where multiple adjectives appear before a noun. They first examined ordering preferences of a sample of 26 adjectives from seven semantic classes (dimension, value, age, physical, shape, color, material) and they subsequently asked participants to determine the naturalness of a canonical/non-canonical adjective pairing (e.g., *the small red chair* versus **the red small chair*). They then conducted a corpus study to determine mean distance from the noun in a series of adjective-adjective-noun phrases that had the adjectives sampled. Finally, they asked participants to rate how subjective each adjective was (e.g., is the adjective *brown* completely objective or completely subjective). Results showed that adjective subjectivity scores accounted for a substantial proportion of the variance in naturalness ratings. Their findings support the hypothesis that adjectives with higher objectivity are located closer to the noun in position than less objective adjectives. Their evidence suggested that modifiers appearing before noun phrase occur in an order that is based in semantics, and adjectives with lower objectivity will be placed further from the noun when stacked against other modifiers.

However, other studies provide evidence that a semantic explanation alone is not the sole factor that controls prenominal adjective ordering. In her multifactorial analysis of adjective order in English, Wulff (2003) claimed that she could predict the order of adjective strings with over 70% accuracy. She stated that previous studies of adjective order did not adequately represent combinations that are seen in spoken language. Her corpus-based analysis examined adjective-adjective-noun combinations (since noun phrases with more than two prenominal adjectives are rarely observed) based on several variables. Her research found that these other factors, such as phonological properties and length of the adjectives used, could affect prenominal adjective order. Her conclusion was that adjective order is not entirely syntactic nor semantic in nature.

Ultimately, it cannot be disputed that English speakers, and speakers of other languages, exhibit a preferred prenominal adjective order. Based on prior research, it is clear that adjective order preferences are not determined by the rules of grammar alone. Rather, adjective placement is more likely to fall within semantic rules or categories. Unfortunately, there has been no fixed method that can predict membership of adjectives into distinct categories, nor is there a method that can predict sequencing when two or more prenominal modifiers occur before the noun. While semantic theories can explain why some adjectives are more closely related than others, the boundaries of these mental classes are unknown, and studies have also shown syntactic influences upon modifier position. As a full explanation cannot be derived from either syntactic or semantic theories, it may be beneficial to turn towards the possibility of a cognitive account of prenominal adjective sequencing.

Relationship between the Visual and Semantic Systems

Prior to discussion of the aims of the current study, it is important to elaborate on properties of the visual system that may influence features that are described by adjectives in language. Both classical and contemporary models of visual search are based in theories of selective visual attention. Subsequently, the visual aspects of cognition that may aid in the organization of the semantic system, along with the implications of this relationship for language generation and processing will be discussed.

In a 2015 paper, authors Moran, Zehetleitner, Liesefled, Muller, and Usher described the classical two-stage model of visual search patterns, which includes parallel processing and serial processing. Parallel processing of visual stimuli occurs in a pre-attentive form, in which more-salient object features are flagged for further processing. In the second stage, called serial processing, attention is focused on these objects. This model assumes that attention is limited; therefore, it follows that certain features must be more likely to modulate visual search patterns than others.

When observing a scene in our world at a given moment in time, the brain must filter stimuli and grant cognitive resources towards the fixation and identification of objects that are the most salient, a process known as selective visual attention. Not all items are deemed equally worthy of visual attention, which is why a person is capable of walking into a pole on the street while texting (a humorous example provided by Wolfe and Horowitz in their 2017 paper), even though both objects were technically visible. This case begs the question: which factors decide patterns of visual attention and visual search? Furthermore, which features of specific objects are most important in processing when selecting objects that warrant attention?

Wolfe and Horowitz (2017) analyzed a series of studies and discovered certain predominant factors that mediate visual search patterns when focusing on a field of objects. The authors argue for the existence of a finite number of factors, given that the brain is unlikely to house a series of attributes by which to guide visual attention. These factors were organized from definite guiding attributes (which have been essentially confirmed by multiple studies), such as color, orientation, and size; to probable guiding attributes, such as luminance polarity, shape, and curvature; to less likely guiding attributes, such as threat level. For Wolfe and Horowitz, research on faces and features of faces have generated divergent conclusions, with evidence that visual search is modulated by expressions of different emotions.

Based on this evidence, it could be predicted that a given person would attend to stimuli experimentally placed within the visual field in the following order: size > luminance polarity and shape > emotional valence. This closely mirrors the intuitive rules for prenominal adjective sequencing which, based on syntactic and semantic theories, would be size > luminance > emotional valence > shape. To this date, no research exists that attempts to determine the potential correlation between features that are most salient in visual search patterns and their corresponding modifiers in language.

Should there be a relationship between the order of features for visual analysis and prenominal adjective ordering, a foundational theory for the genesis of the semantic system could be developed. Research has shown that the concepts that humans communicate with language rely on the access and integration of our varied experiences (Binder & Desai, 2011). This information is stored in several brain regions, with greater activation noted in the left hemisphere, an area which is also essential for language function in a vast majority of people (Binder, Desai, Graves, & Conant, 2009). Semantic memory is thought to be organized according to attributes, with certain brain regions that are specialized for retrieval of these concepts. The results of a study attempting to determine whether visual access and syntactic constraints on the position of words in sentences could provide evidence that language structures are determined largely by properties of the semantic system.

Aims of Present Study

The evidence cited above has shown that neither syntactic nor semantic theories of prenominal adjective ordering alone are sufficient to explain the phenomenon of preferred ordering of multiple adjectives placed before a noun in English. Further research indicates that semantic and cognitive domains interact with each other in the processing of both linguistic and visual stimuli. The focus of the current study is to determine whether prenominal adjective ordering may have a cognitive-semantic basis, in that certain features that correspond to adjectives—such as words describing size or luminance of an object—may be more visually salient in visual perception. Its purpose is to determine whether prenominal adjective order in English is highly correlated with visual search patterns. No current studies have investigated the possibility that human language is organized according to which visual features are easiest to access and process. Since it is clear that the domains of vision and semantics interact, it could be that prenominal adjective order is driven by properties of the visual system.

Therefore, the aims of the current study are threefold. The first goal is to determine whether individuals use the same scan pattern all the time, or whether scan pattern depends on image type. It is hypothesized that visual search pattern—whether orthographic, clockwise, or random—will not vary with the type of image presented in the visual field during an oddball paradigm (examples of these search patterns are shown in Figure 1). The second purpose is to examine visual search patterns of English-speaking individuals, and to determine whether their scan patterns follow an order that is correlated with prenominal adjective order. It is hypothesized that participants will follow a visual search pattern that is correlated with prenominal adjective order (size > luminance > emotional valence > geometry/shape). The third purpose is to determine whether inversions of certain adjective classes, measured through acceptability ratings of canonical and non-canonical adjective pairs, are deemed more acceptable than other inversions (i.e., are violations of the size-geometry order more or less acceptable than violations of the valence-geometry order). It is hypothesized that violations of adjective pairs that are typically placed further from the noun will differ significantly on acceptability ratings when compared to adjectives typically placed closer to the noun. The fourth and final purpose of the study is to describe correlations between the cognitive and semantic systems. Data drawn from these correlations may yield information about the organization of semantic concepts and language in the brain, and could influence ideas of whether language influences thought or vice versa.

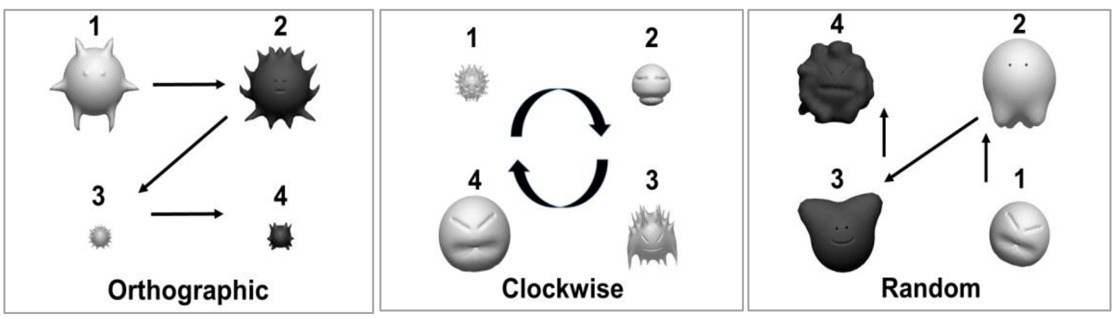

Figure 1. Systematic Search Patterns Analyzed in Experiments 1 and 2

CHAPTER 2

METHODS

To determine whether prenominal adjective order has its basis in visual perception, we presented arrays of four stimulus items, lovingly referred to as “blobs,” to healthy young adults with the intent of recording the path of visual search (i.e., scan path). We hypothesized that adjectives would correspond to visual features that would be more salient in human perception, which may form the basis for prenominal adjective order in various human language forms. A survey was also designed to determine whether inversions of certain adjective categories were more or less acceptable than inversions of other adjective categories. In the following section, participant demographics, inclusion and exclusion criteria, stimulus design, equipment, and the three experimental designs will be described in detail.

Participants

Twenty healthy adults (N = 20) were recruited from Temple University (mean age = 24.1, range = 19-53, SD = 7.44). All participants were native speakers of English who had completed a secondary level of education (mean = 15.95 years, range = 13-22, SD = 2.09) who reported that they were in good neurological health. Visual acuity of all participants was evaluated using a Snellen eye chart, with corrected vision of 20/30 as the criteria for inclusion. Prior to the experiment participants gave informed consent, and were asked to remove hard contact lenses, glare resistant eyeglasses, and eye makeup prior to initiation of the eye tracking protocol.

Stimuli

The stimuli used in the experiment were designed to resemble inanimate and animate objects (blobs and creatures) in the 3-D sculpting program Sculptris (Pixologic Inc., Los Angeles, CA) The included stimuli were normed using a 5-point Likert-Scale ratings of four image categories—size, luminance, emotional valence, and geometry on Qualtrics survey software (Qualtrics Inc., Provo UT). Two binary characteristics were examined within each image category for size (big/small), luminance (bright/dark), emotional valence (cute/scary), and geometry (smooth/spiky), adding to a total of 8 experimental features. Average ratings for each object in each of the four characteristics were calculated with a cutoff of 1.5 standard deviations from the mean as criteria for inclusion as stimulus items (mean = 3). In experiments 2 and 3, the figures were arranged in 2 x 2 arrays with each stimulus centered within the four quadrants and located equidistant from a central point on a 10” by 7.5” slide created with PowerPoint software.

Eye Tracking Equipment and Protocols

Participants were seated at a desk in a quiet, windowless room under constant fluorescent lighting. Participants were positioned on an optician’s chin rest to maintain head stability for calibration and for the duration of the eye tracking experiments. Stimuli were presented on a 17-inch LCD monitor positioned 60 cm from the chin rest. A remote 120 Hz infrared eye tracking sensor was placed at the base of the monitor (SMI Inc. Red-M). Following positioning, the participant underwent a 5-point calibration and validation procedure, in which the participant optically tracked a black disc moving across a white screen. The experimental protocols were initiated when calibration was determined to be <0.5 degrees. Stimuli delivery and data collection was accomplished with Experiment Center Software (SMI Inc., Boston MA) and analyzed using BeGaze (SMI Inc., Boston MA).

Experimental Procedures

Experiment 1: Oddball Paradigm

An experimental protocol was designed to determine whether participants’ eye movements exhibited a tendency to follow a single, predictable path across a visual field with multiple stimuli regardless of image category. The 36 experimental slides were arranged in 2 x 2 stimulus arrays of four stimuli that were randomly assigned to either the upper right (UR), upper left (UL), lower right (LR), or lower left (LL) corners of the screen.

Participants were seated at the eye tracker and completed steps necessary for calibration and validation. They subsequently received training via written directions accompanied by aural instructions from the examiner prior to viewing the experimental slides. The stimuli represented on each slide all belonged to one of three general visual categories: numbers, letters, and shapes (n = 12 per category). Of the four stimuli presented in each array, participants viewed three stimuli with the same shape (spiky/smooth, filled/hollow, square/round), alphabetic feature (consonant/vowel, letter sequence), or numerical feature (even/odd, double/random, numerical sequence), in addition to one deviant stimulus item. Participants were instructed to visually identify the “oddball” as quickly as possible and to focus their gaze upon the selected target until the next slide appeared. Presentation of the subsequent slide was triggered after 2000 milliseconds. A fixation point appeared in the middle of the screen between trials.

The participants’ gaze and scan path were recorded on a SensoMotoric Instruments (SMI) eye tracker using the program’s data recording and analysis programs: Experiment Center and BeGaze. Reaction time, accuracy, and scan path were recorded and analyzed for recurring (directional) patterns across image type.

Experiment 2: Looking Bias

A second protocol was designed to determine whether participants’ gaze was drawn to features in an order comparable to the order observed in the grammatical norms that govern prenominal adjective ordering in spoken language (e.g., the “big bright cute smooth one”). In other words, are features furthest from the noun (e.g., size) more salient in visual perception than words closer to the noun (e.g., smooth). This research question was examined by monitoring participants’ eye movements as they viewed the stimuli during the experimental task.

It was hypothesized that, of the four image categories, size would be the most visually-salient feature, followed by luminance, emotional valence, and finally geometry. Therefore, a total of 15 trials were designed for each of the six experimental conditions (size > luminance, size > valence, size > geometry; luminance > valence, luminance > geometry; valence > geometry) totaling 60 experimental trials. Since each “blob” was normed on all four image features, we controlled two image categories at a time, and determined via sequencing patterns which of the tested features was most salient. For example, in a given array testing differences in salience between luminance and valence, all four stimuli were controlled for the remaining two categories (i.e., all stimuli were the same size and geometric shape). In this situation, the stimuli would differ on their luminance (with two bright and two dark objects) and their valence (with two cute and two scary objects). To determine whether participants viewed size or valence as the most salient feature, scan path sequence data was analyzed to detect whether search focused on bright/dark objects before cute/scary objects or vice versa.

Each of the four categories was combined in all possible combinations to determine the order of the most engaging categories in descending order. The target stimuli were assigned to a random area of interest (AOI) on a 2x2 array with each stimulus centered within equal quadrants and with each stimulus item located equidistant from a central point on a 10” by 7.5” slide created with PowerPoint software. To maintain experimental validity and to prevent initial unrelated bias from causing participants to look at certain images over others, the protocol was disguised as a 2-back memory task. Each experimental trial consisted of 3000 milliseconds of a visual array, during which the participants were instructed to attempt to recall each image and its location as accurately as possible. A detailed diagram of the structure and duration of the trials is provided in Figure 2.

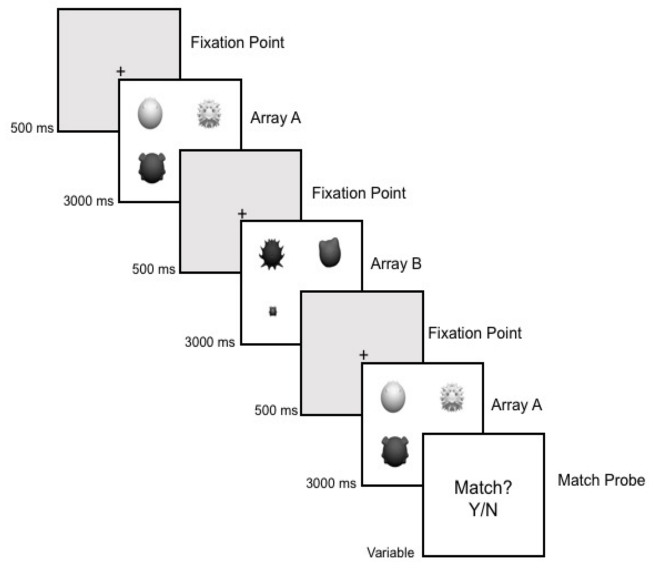

Figure 2. Looking Bias Trial Structure and Durations

A total of 21 probe trials were quasi-randomly interspersed throughout the 60 experimental arrays, during which the participant was asked to indicate on a keypad (*Yes* or *No*) whether or not the slide perfectly matched the array that was presented 2 slides previously. Trial and probe arrays were both preceded and followed by 500 milliseconds of a fixation point located in the center of the screen. The order in which participants gazed at each area of interest (AOI), specifically the four quadrants containing the stimuli, was analyzed using BeGaze software (SMI Inc., Boston, MA). Testing all possible pairs/combinations allowed us to analyze the expression of looking bias.

Experiment 3: Canonical & Non-canonical Acceptability

Immediately following the second experiment, participants completed a survey with Qualtrics software (Qualtrics Inc., Provo UT) whereby they provided acceptability ratings of both canonical (e.g., the *creepy black* one) and non-canonical (e.g., the *pointy tiny* one) adjective pairings. With a 7-point Likert Scale, participants rated a total of 96 pairwise comparisons of adjectives from each of the 4 experimental categories (size, luminance, emotional valence, geometry) presented in both canonical and non-canonical orders.

The survey examined the adjectives that correlated with the 8 binary feature categories examined with eye tracking in experiments 2 and 3 (big/small, bright/dark, cute/scary, smooth/spiky). An adjective pool was created with 12 synonyms per binary within each category (for example, miniature, tiny, and miniscule were used as synonyms for small). The 24 adjectives within the four domains were randomized and combined in both canonical and non-canonical variations according to the experimental condition (e.g., it was predicted that a combination wherein an adjective for size preceded any other adjective category would be more acceptable to participants on average). We examined a total of 6 experimental conditions (size > valence, size > geometry, size > luminance, luminance > valence, luminance > geometry, valence > geometry) with 8 adjectives used per condition. There were 48 canonical and 48 non-canonical pairs, adding to a total of 96 experimental trials. These trials were randomized and presented to participants, who rated them. A score of 1 was least acceptable, a score of 7 was most acceptable.

CHAPTER 3

RESULTS

Experiment 1: Oddball Paradigm

Each participant's eye tracking data was exported and analyzed for scan sequence to determine which objects commanded the most interest. Scan path information was also analyzed to determine whether participants performed a single type of search (e.g., orthographic, clockwise, or random) across the three image categories (numbers, letters, and shapes).

First, we analyzed the proportion of each scan pattern that was used during the 36 trials within each individual participant. Raw data are outlined in Table 2, with group analysis located in Table 3. Based on evidence recorded from within-subject scan data showed in Table 1 and subsequent hypothesis testing, there is insufficient evidence that any participant uses a single scan pattern to view 2x2 arrays of varied objects across the 36 experimental trials at a 95% confidence level.

Table 2. Scan Pattern Proportion Observed

Subject	Orthographic	Clockwise	Random	No Fixation
1	0.17	0.36	0.22	0.25
2	-	0.22	0.28	0.50
3	0.11	0.36	0.22	0.31
4	-	0.28	0.14	0.58
5	0.22	0.39	0.03	0.36
6	0.47	0.14	0.08	0.31
7	0.03	0.25	0.17	0.56
8	0.08	0.08	0.25	0.58
9	0.25	0.56	0.03	0.17
10	0.08	-	0.50	0.42
11	0.03	0.64	0.03	0.31
12	0.08	0.56	0.03	0.33
13	0.50	0.25	-	0.25
14	0.03	0.22	0.36	0.39
15	0.17	0.42	0.42	-
16	0.11	0.22	0.47	0.19
17	0.22	0.25	0.17	0.36
18	-	0.36	0.14	0.50
19	0.08	0.22	0.33	0.36
20	0.33	0.25	0.06	0.36

Note. Values correspond with the frequency of use of each systematic pattern within participants. Higher proportions indicate greater likelihood to use the selected systematic pattern when viewing different image arrays (composed of letters, numbers, and shapes).

Table 2 demonstrates that across participants, when viewing stimulus arrays of various image types, there is a statically significant propensity to use a clockwise scan pattern in lieu of an orthographic or random scan pattern at a 95% confidence level. In fact, there is a slightly higher propensity to view objects without fixation, followed by a clockwise pattern. An orthographic pattern was used with the lowest frequency observed. The high standard deviation across scan patterns compared with their respective means would indicate that there is high variability in preferred scan pattern across the twenty participants.

Table 3. Scan Pattern Proportion Observed Across Subjects

Measure	Orthographic	Clockwise	Random	No Fixation
Average	0.15	0.30	0.20	0.35
SD	0.15	0.16	0.16	0.15

Note. Values indicate average proportion of scan pattern use across all individuals and all trials and standard deviation from the mean.

A possible explanation is that different image categories require or result in specific types of visual search patterns for identification. Subsequent analyses attempted to discern whether differences in scan patterns caused by pattern preferences for specific mage types (e.g., an orthographic pattern is preferred when viewing letters versus a clockwise pattern for numbers) could cause the high variability observed in across-participant scan path data. In Table 4, results across image types are summarized. The data was analyzed using a two-tailed hypothesis test for proportions at a confidence level of 95%.

Table 4. Scan Pattern Proportion Within Image Type

Category Type	Orthographic	Clockwise	Random	No Fixation
Letters	0.24	0.31	0.23	0.22
Numbers	0.19	**0.45**	0.22	0.15
Shapes	0.03	0.18	0.13	**0.67**

Note: A two-tailed hypothesis test for proportions was performed with a confidence level of 95%. Statistically significant contrasts are highlighted in bold.

The shaded cells in Table 4 show systematically higher propensity for certain pattern types. Specifically, participants were significantly more likely to use a clockwise pattern when viewing numbers. Participants were unlikely to use an orthographic pattern when observing shapes, with a marked tendency to not fixate at all upon images within the stimulus arrays. This result could be easily explained due to the strong apparent contrast between the target and foils when viewing shapes. For letters, there is a tendency towards the use of a clockwise pattern, however there is not sufficient evidence to conclude any pattern was preferred by participants. Results are represented graphically below in Figure 3.

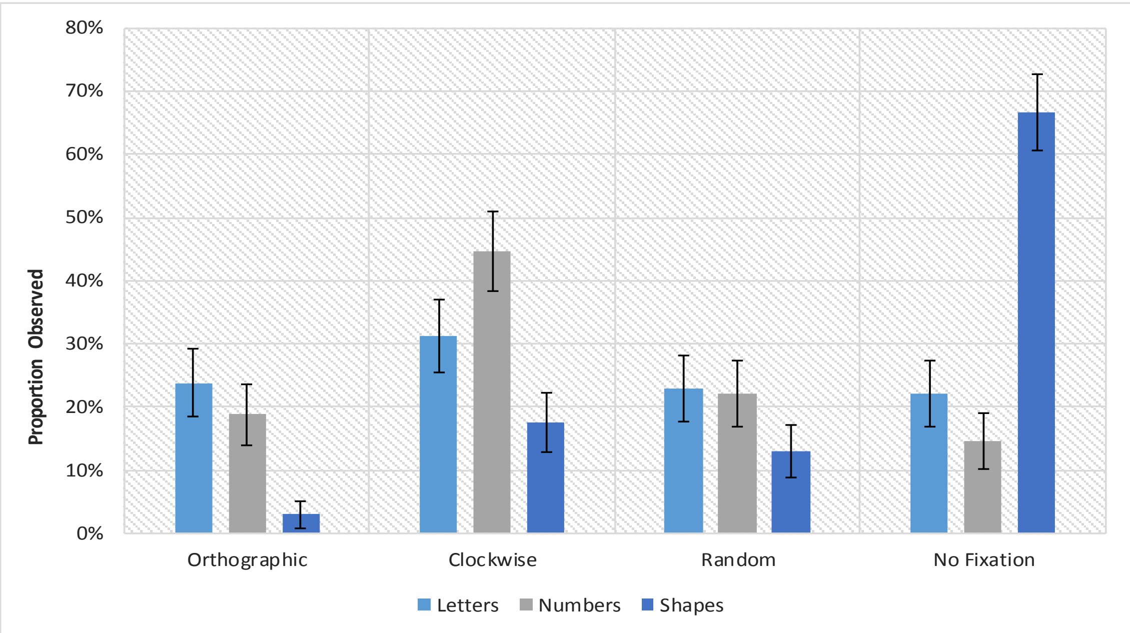

Figure 3. Scan Pattern Observed by Image Type
Note. Error bars indicate a 95% confidence level of proportions across all subjects and trials.

Experiment 2: Looking Bias

Following completion of the protocol, participants' scan path data were exported and analyzed for recurring patterns within and across trials and subjects. To determine whether specific image traits had greater visual salience, the sequence in which AOIs were viewed was recorded. The specific scan sequence was used to analyze whether participants looked at objects in an order that corresponded to canonical rules for prenominal adjective order in English (e.g., they looked at big things before bright things, or cute things before spiky things when all other traits were similar across all four stimuli), whether they looked at objects in an order that corresponded to an inversion of canonical rules according to rules for prenominal adjective sequencing (e.g., they looked at spiky things before big things, or bright things before small things), or whether there was no discernable pattern. We created 10 trials per matchup (total N = 60), with an equal number of matchups between the 4 categories tested (size, luminance, valence, and geometry) to enable the selection of an ordinal ranking system to compare the salience of each.

Table 5. Ordinal Ranking Data Across Binary Image Categories

Subject	**Size**	**Luminance**	**Valence**	**Geometry**	**No Fixation**
1	10	8	7	9	24
2	9	14	8	9	18
3	10	14	7	10	17
4	11	10	9	8	20
5	13	15	7	8	15
6	17	12	12	8	9
7	12	8	8	9	21
8	10	11	9	11	17
9	12	15	7	9	15
10	12	11	6	10	19
11	11	13	8	8	18
12	12	16	10	8	12
13	13	10	10	10	15
14	16	10	9	9	14
15	14	10	9	9	16
16	12	6	3	11	26
17	11	17	7	10	13
18	14	9	10	8	17
19	10	11	6	8	23
20	11	8	13	9	17
Total	240	228	165	181	346
Average	12	11.4	8.25	9.05	17.3
SD	2.05	3.02	2.22	1.00	4.12

Note. Values indicate the cumulative number of wins obtained from each tested matchup (Size-Luminance, Size-Valence, Size-Geometry, Luminance-Valence, Luminance-Geometry, and Valence-Geometry). The number of wins was then used to perform a statistical test of ordinal rank. Average and standard deviation were calculated.

When looking at ordinal ranking data for individual subjects (summarized in Table 5), it was found that participants looked at large/small stimuli first with the greatest frequency when compared to other binary visual categories, followed by bright/dark stimuli, smooth/spiky stimuli, and cute/scary stimuli respectively. This is not in alignment with the predicted ordinal ranking attained by prenominal adjective sequencing rules (big/small > bright/dark > cute/scary > smooth/spiky). These results are represented visually in Figure 4. While results trended towards the predicted order of visual saliency that aligns with prenominal adjective sequencing (with the exception of geometry being viewed more frequently over valence), the trend was not statistically significant.

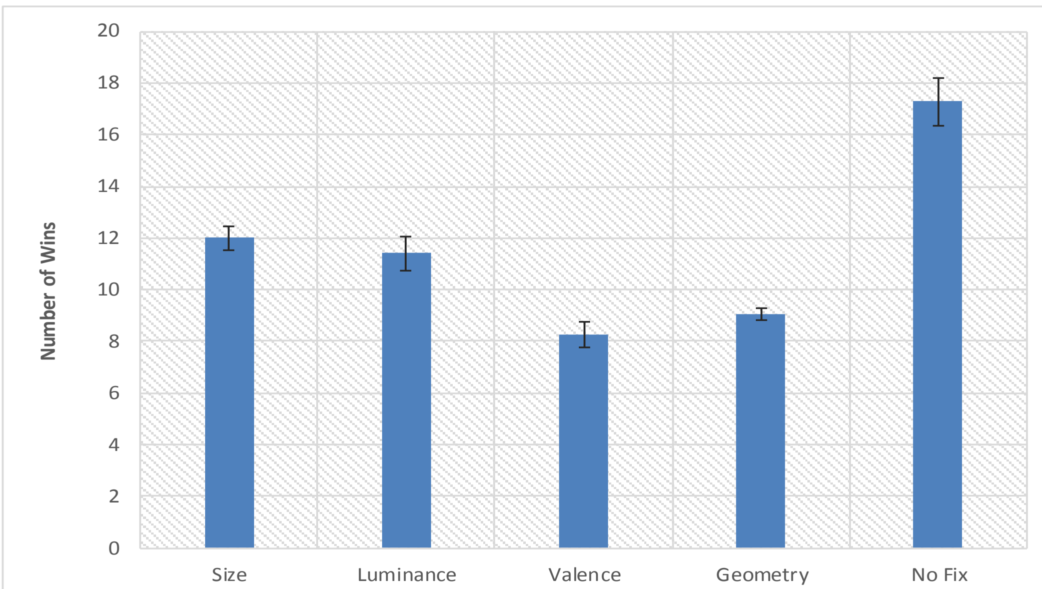

Figure 4. Ordinal Ranking Data by Binary Feature Category

Note. Error bars represent a 95% confidence level of expected wins in an ordinal rank comparison of each of the four image categories.

Further testing was performed in order to determine whether statistically significant differences could be noted between individual category matchups tested within the experiment. A one-tail hypothesis test measuring the difference in proportions in individual matchups was performed with a 95% confidence level. This test was used to determine the likelihood of one trait being more visually salient than another tested trait. Results are summarized in Table 6, with shaded items indicating statistically significant contrasts. This data set only includes trials where fixations were recorded.

Table 6 . Ordinal Ranking Data Proportions

Pairwise Comparison	Size	Luminance	Valence	Geometry
Size	**-**	0.49	**0.75**	0.53
Luminance	0.51	**-**	0.45	**0.59**
Valence	0.25	0.55	**-**	**0.58**
Geometry	0.47	0.41	0.42	**-**

Note: Statistically significant contrasts are highlighted in bold.

As seen in the table above, size is substantially more salient to the visual field than valence. Luminance and valence were also significantly more salient than geometry. However, in all other cases, no contrast could be determined. In comparisons of size and geometry, there was a trend towards size as a more salient feature, but the data collected were not significant. Additionally, there was a strong trend towards valence as a more salient feature than luminance, however these results were also not statistically significant.

Experiment 3: Canonical and Non-Canonical Acceptability

Following completion of the Qualtrics survey, participant data (collected in the form of a Likert Scale rating between 1 and 7, with 7 being very acceptable) was exported to a spreadsheet where average acceptability ratings for each participant were calculated across each binary category tested (e.g., size-luminance, valence-geometry, etc.). The overall acceptability of canonical versus non-canonical adjective pairings was measured using a one-tailed t-test with a 95% confidence level. The data indicate that canonical pairs were more preferred by participants overall (μ canonical = 4.81, μ non-canonical = 4.54, $t = 2.84$, $p = 0.0022$).

Subsequently, the acceptability of canonical and non-canonical versions of each individual adjective pair were contrasted using a one-tailed hypothesis test for the difference in means using a 95% confidence level. Results of this test are displayed in Figure 5, which shows that in every category tested the canonical pair was preferred, even though the canonical pairs were not significantly more preferred in all matchups. A one-tailed test was performed to confirm a preference for canonical over non-canonical adjective pairs.

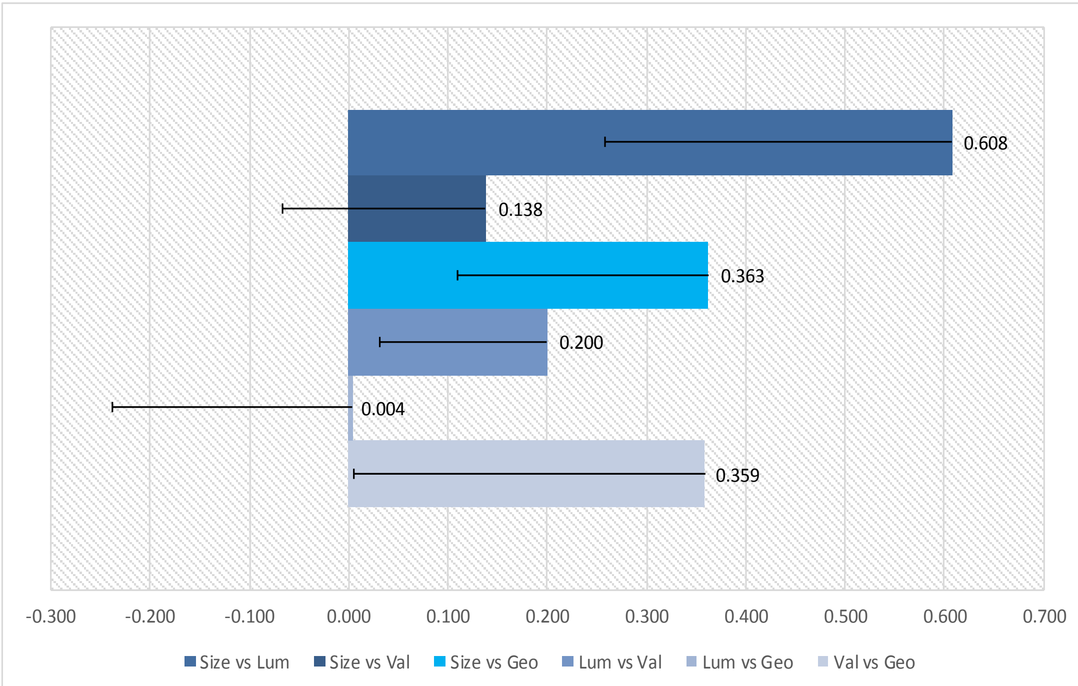

Figure 5. Differences in Acceptability of Canonical Versus Non-Canonical Adjective Pairs

Note. Data show the difference between mean acceptability rating given across all individuals for a given matchup in their canonical and non-canonical order, A positive value demonstrates the degree to which canonical word pairs were preferred, with a negative value demonstrating the possibility that a non-canonical ordering was preferred. Error bars show a 95% confidence level for the minimum acceptability rating.

The largest difference was observed in the size-luminance comparison, with the effect size of the canonical version calculated as 12%. Similar analyses were performed within the size-geometry, luminance-valence, and valence-geometry comparisons, which demonstrated that canonical versions were preferred in all categories. While participants preferred adjective pairs where size adjectives placed before adjectives describing emotional valence, this preference was not statistically significant. A similar effect was observed in acceptability ratings between adjectives that describe luminance and adjectives that describe geometric properties.

A second t-test was performed to compare acceptability of each type of non-canonical adjective pairing and to determine whether inversions of adjective pairs that are typically located closer to the noun in position (e.g., adjectives comparing shape) were more acceptable than inversions of adjective pairs located farther from the noun (e.g., adjectives comparing size). Results are exhibited in Table 7. Following a t-test comparing all non-canonical adjective pairs (Luminance-Size, Valence-Size, Geometry-Size; Valence-Luminance, Geometry-Luminance; Geometry-Valence) with 95% confidence, it was determined that only inversions of the Size-Luminance pairing were significantly less acceptable than non-canonical pairs overall.

Table 7. Acceptability of Non-Canonical Adjective Pairs

Feature Matchup	Mean	t	p-value
Overall Non-Canonical	4.54	-	-
Luminance-Size	**4.34**	**2.21**	**0.02**
Valence-Size	4.62	-0.95	0.83
Geometry-Size	4.63	-1.01	0.84
Valence-Luminance	4.46	0.91	0.18
Geometry-Luminance	4.62	-0.97	0.83
Geometry-Valence	4.55	-0.16	0.56

Note: Statistically significant contrasts are highlighted in bold

CHAPTER 4

DISCUSSION

Overall results were generally in line with each of the three hypotheses presented. Participants' use of systematic search patterns did not vary significantly when they were presented with shapes. Participants viewed objects in an order that was correlated with prenominal adjective ordering with varying levels of significance. Finally, acceptability judgments indicated that inversions of adjective classes that are typically placed closer to the noun were generally more acceptable than inversions of adjective classes typically placed further from the noun. Below is a discussion of the results obtained from each individual experiment, as well as their implications for the overall hypothesis that the rules which govern the position of multiple adjectives stacked before the noun may be derived from the order in which specific features of objects in the visual field are processed.

The purpose of the first experiment was to determine whether individuals use the same scan pattern all the time, or whether scan pattern depends on image type. It was hypothesized that scan patterns would not vary with the type of image presented in the visual field. Results indicate that participants were significantly more likely to use a clockwise search pattern when viewing an array of four numbers. However, this was the only scan pattern of significance across letters, numbers, and shapes. Regardless, the first hypothesis was disproved. Upon deeper analysis of the results, it was found that participants were significantly less likely to fixate on the stimuli when the stimuli were comprised of shapes, which could have accounted for the high level of variability between visual search patterns when presented with arrays of stimuli.

The purpose of the second experiment was to examine the visual search patterns of English-speaking individuals, and to determine whether their scan patterns follow an order that is correlated with prenominal adjective order. It was hypothesized that participants would follow a visual search pattern that is correlated with prenominal adjective order in English (i.e., the features most salient to English speakers would be correlated with language structure in that they would view objects in terms of size > luminance > emotional valence > geometry/shape). Results showed that, while feature salience was correlated with the expected order based on language rules, the tendencies observed were not statistically significant. While size was more salient than valence, and geometry was the least salient visual feature overall, other trait matchups either were not correlated with prenominal adjective order in English, or there was no significant difference between effect sizes. One possible explanation for the inconsistencies in the data was the experimental design, with participants being instructed to attempt to recall each of the images in a modified n-back memory task. The format of the task in the second experiment could have increased the incidence of systematic scan patterns (e.g., clockwise) during the protocol. Another inherent issue in data analysis was the considerable number of trials that could not be analyzed due to a lack of fixations. These results could be explained by the results from Experiment 1, which showed that participants were significantly more likely to not fixate on stimuli when they belonged to the image category of shapes, however, the lack of fixations in Experiment 2 could have skewed results.

In regard to the second experiment, future studies could include more categories of adjective classes. Previous experiments have suggested that stimuli with facial features yield inconsistent results. With the exception of the switch between valence and geometry image categories, the results attained do trend towards alignment with prenominal adjective ordering rules, which supports the hypothesis that prenominal adjective ordering can be derived from the semantic system via feature salience as determined by the visual system. Other categories could determine whether this trend is significant, or whether humans are predisposed to searching for categories in a more random order. Subsequent studies could also analyze whether certain features within the categories examined (ex. big or bright objects) are more visually salient than all others (ex. small or dark objects).

The purpose of the third experiment was to determine whether inversions of certain adjective classes, measured through acceptability ratings of canonical and non-canonical adjective pairs, would be deemed more acceptable than other inversions. Specifically, are inversions of adjectives that are typically placed further from the noun in position (such as adjectives of size) more unacceptable than inversions of adjectives placed closer to the noun (such as adjectives of emotional valence). It was hypothesized that violations of adjective pairs that are typically placed further from the noun will differ significantly on acceptability ratings when compared to adjectives typically placed closer to the noun. Results show that canonical adjective pairs were preferred over non-canonical pairs with statistical significance. Additionally, of the canonical versions of the tested adjective classes, it was determined that canonical orderings were preferred with statistical significance in the size-luminance, size-geometry, luminance-valence, and valence-geometry matchups. A trend without significance were noted for the canonical versions of size-valence adjective pairs. In the case of the luminance-geometry adjective pairings, the data indicate that participants had no preference between canonical and non-canonical pairings.

Where preferences could not be identified, as in the case of size-valence and luminance-geometry, there are several possible explanations. While the absence of preference is possible, it is unlikely given previous studies and grammatical rules. Another explanation is that the types of words selected for use in the Qualtrics survey were unfamiliar, less frequent, and resulted in decreased ratings overall (ex., the theoretically canonical combination of *The big inky one* may automatically lead to decreased acceptability). Past studies have indicated the possibility of adjective order being influenced by properties of the noun being modified (see examples from Martin, 1969). Further research could be performed to determine whether acceptability ratings of these pairs change based on the noun (e.g., *the shadowy soft dog* versus *the shadowy soft cloak*).

Upon analysis of non-canonical adjective pairings to determine whether certain adjective inversions were less acceptable than others, it was found that only one matchup (size-luminance) was significantly more unacceptable than other non-canonical adjective pairings studied. This would indicate that, while canonical adjective pairs are preferred over non-canonical pairs, there is insufficient evidence from this study to determine the relationship between the position of the targeted adjective classes in relation to the noun and the perceived acceptability of a given prenominal adjective sequence. It should be noted that this study required participants to read 96 predetermined adjective pairs. A future study could ask participants to generate an adjective sequence independently when given the feature categories (i.e., write a sentence to describe an object in terms of its size, brightness, valence, and geometry).

Conclusions

This investigation provided preliminary evidence to indicate that visual search patterns may influence order preferences when multiple adjectives are placed before a noun in English. While there were variations in significance, the results of the three experiments generally showed trends indicating that systematic visual search patterns do not vary based on the type of image presented, that visual scan patterns are correlated with prenominal adjective ordering rules, and that inversions of adjective classes typically placed further from the noun are less acceptable than inversions of adjective classes placed closer from the noun. Further studies, which may target the generation of multiple adjectives to describe a single object, or may determine whether abstract adjectives fall into a hierarchy based on a defining category (e.g., morality, frequency, or concreteness), could be performed to advance discussion on the relationship between grammatical rules in language and aspects of cognition or the semantic system.

REFERENCES

All Things Considered. BBC Editor Highlights Often Overlooked English Language Rule. (2016, September, 08). Retrieved March 1, 2017, from http://www.npr.org/2016/09/08/493157931/bbc-editor-highlights-often-overlooked-english-language-rule

Annear, S. S. (1964). The Ordering of Pre-nominal Modifiers in English.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767-2796.

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in cognitive sciences*, *15*(11), 527-536.

British Counsel: Adjective Order. (n.d.). Retrieved from https://learnenglish.britishcouncil.org/en/english-grammar/adjectives

Cambridge Dictionary: Adjective Order (n.d.). Retrieved from http://dictionary.cambridge.org/us/grammar/british-grammar/about-adjectives-and-adverbs/adjectives-order

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on information theory*, *2*(3), 113-124.

Chomsky, N. (1959). On certain formal properties of grammars. *Information and control*, *2*(2), 137-167.

Cinque, G. (1994). On the evidence for partial N-movement in the Romance DP. *Paths towards universal grammar*, 85-110.

Danks, J. H., & Glucksberg, S. (1971). Psychological scaling of adjective orders. *Journal of Verbal Learning and Verbal Behavior*, *10*(1), 63-67.

Danks, J. H., & Schwenk, M. A. (1972). Prenominal adjective order and communication context. *Journal of verbal learning and verbal behavior*, *11*(2), 183-187.

Education First: Ordering Multiple Adjectives. (n.d.). Retrieved from http://www.ef.edu/english-resources/english-grammar/ordering-multiple-adjectives/

Forsyth, M. (2013). *The elements of eloquence: How to turn the perfect English phrase*. Icon Books Ltd.

Martin, J. E. (1969). Semantic determinants of preferred adjective order.*Journal of Verbal Learning and Verbal Behavior*, *8*(6), 697-704.

Moran, R., Zehetleitner, M., Liesefeld, H. R., Müller, H. J., & Usher, M. (2016). Serial vs. parallel models of attention in visual search: accounting for benchmark RT-distributions. *Psychonomic bulletin & review*, *23*(5), 1300-1315.

Scontras, G., Degen, J., & Goodman, N. D. (2017) Subjectivity Predicts Adjective Ordering Preferences. *at Open Mind*: *Discoveries In Cognitive Science.* Advance Publication. doi:10/11.62/opmi_a_00005.

Sussex, R. (1974). The deep structure of adjectives in noun phrases. *Journal of Linguistics*, *10*(01), 111-131.

University of Victoria English Language Centre Study Zone: Adjective Order. (2011). Retrieved from http://web2.uvcs.uvic.ca/elc/studyzone/410/grammar/adjord.htm

Vendler, Z. (1967). Linguistics in philosophy.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*, 0058.

Wulff, S. (2003). A multifactorial corpus analysis of adjective order in English.*International Journal of Corpus Linguistics*, *8*(2), 245-282.

APPENDIX A

QUALTRICS SURVEY STRUCTURE

In this part of the survey, you will be asked to rate the acceptability of a series of sentences. A score of 1 indicates that the sentence is very unacceptable. You would never hear the sentence spoken by a native speaker of English. A score of 7 indicates that the sentence is very acceptable. You are likely to hear this sentence spoken by a native speaker of English.

On a scale from 1-7, how acceptable is the following sentence:

1- Very unacceptable
2- Unacceptable
3- Somewhat unacceptable
4- Neutral
5- Somewhat acceptable
6- Acceptable
7- Very acceptable

Survey item format: Look at the $TRAIT_1$ $TRAIT_2$ one.

APPENDIX B

EXPERIMENT 3: SUMMARY OF PRESENTED STIMULI

Table B1. Adjective Pairs Presented in the Qualtrics Survey

Adjective Pair	Canonical	Trait 1	Trait 2	Prediction
gargantuan pitch-black	1	Size	Lum	Size>Lum
enormous white	1	Size	Lum	Size>Lum
minuscule shadowy	1	Size	Lum	Size>Lum
big inky	1	Size	Lum	Size>Lum
massive ebony	1	Size	Lum	Size>Lum
mini unlit	1	Size	Lum	Size>Lum
little pale	1	Size	Lum	Size>Lum
undersized underlit	1	Size	Lum	Size>Lum
pitch-black gargantuan	0	Size	Lum	Size>Lum
pale little	0	Size	Lum	Size>Lum
shadowy miniscule	0	Size	Lum	Size>Lum
unlit mini	0	Size	Lum	Size>Lum
white enormous	0	Size	Lum	Size>Lum
inky big	0	Size	Lum	Size>Lum
underlit undersized	0	Size	Lum	Size>Lum
ebony massive	0	Size	Lum	Size>Lum
miniature fearsome	1	Size	Val	Size>Val
gigantic bone-chilling	1	Size	Val	Size>Val
teeny-weeny gorgeous	1	Size	Val	Size>Val
immense cute	1	Size	Val	Size>Val
giant unnerving	1	Size	Val	Size>Val
large horrifying	1	Size	Val	Size>Val
colossal delightful	1	Size	Val	Size>Val
huge attractive	1	Size	Val	Size>Val
unnerving giant	0	Size	Val	Size>Val
gorgeous teeny-weeny	0	Size	Val	Size>Val
horrifying large	0	Size	Val	Size>Val
attractive huge	0	Size	Val	Size>Val
delightful colossal	0	Size	Val	Size>Val
bone-chilling gigantic	0	Size	Val	Size>Val
cute immense	0	Size	Val	Size>Val
fearsome miniature	0	Size	Val	Size>Val

Note. Canonical pairs were determined by English adjective ordering rules and are indicated by a value of 1, while non-canonical pairs are shaded in red and correspond to a value of zero.

Table B1 (cont.). Adjective Pairs Presented in the Qualtrics Survey

Adjective Pair	Canonical	Trait 1	Trait 2	Prediction
small rugged	1	Size	Geo	Size>Geo
oversized bumpy	1	Size	Geo	Size>Geo
humongous glassy	1	Size	Geo	Size>Geo
teeny polished	1	Size	Geo	Size>Geo
compact jagged	1	Size	Geo	Size>Geo
tiny pointy	1	Size	Geo	Size>Geo
minute coarse	1	Size	Geo	Size>Geo
itty-bitty smooth	1	Size	Geo	Size>Geo
jagged compact	0	Size	Geo	Size>Geo
rugged small	0	Size	Geo	Size>Geo
pointy tiny	0	Size	Geo	Size>Geo
glassy humongous	0	Size	Geo	Size>Geo
bumpy oversized	0	Size	Geo	Size>Geo
polished teeny	0	Size	Geo	Size>Geo
coarse minute	0	Size	Geo	Size>Geo
smooth itty-bitty	0	Size	Geo	Size>Geo
sweet dark	1	Lum	Val	Lum>Val
eerie luminous	1	Lum	Val	Lum>Val
terrifying fluorescent	1	Lum	Val	Lum>Val
darling dusky	1	Lum	Val	Lum>Val
creepy black	1	Lum	Val	Lum>Val
scary beaming	1	Lum	Val	Lum>Val
sinister dazzling	1	Lum	Val	Lum>Val
lovable sparkling	1	Lum	Val	Lum>Val
fluorescent terrifying	0	Lum	Val	Lum>Val
luminous eerie	0	Lum	Val	Lum>Val
beaming scary	0	Lum	Val	Lum>Val
dazzling sinister	0	Lum	Val	Lum>Val
black creepy	0	Lum	Val	Lum>Val
dark sweet	0	Lum	Val	Lum>Val
dusky darling	0	Lum	Val	Lum>Val
sparkling lovable	0	Lum	Val	Lum>Val

Note. Canonical pairs were determined by English adjective ordering rules and are indicated by a value of 1, while non-canonical pairs are shaded in red and correspond to a value of zero.

Table B1 (cont.). Adjective Pairs Presented in the Qualtrics Survey

Adjective Pair	Canonical	Trait 1	Trait 2	Prediction
slick glowing	1	Lum	Geo	Lum>Geo
velvety shady	1	Lum	Geo	Lum>Geo
circular unilluminated	1	Lum	Geo	Lum>Geo
uneven light	1	Lum	Geo	Lum>Geo
choppy jet-black	1	Lum	Geo	Lum>Geo
sleek radiant	1	Lum	Geo	Lum>Geo
flat gleaming	1	Lum	Geo	Lum>Geo
spiny bright	1	Lum	Geo	Lum>Geo
light uneven	0	Lum	Geo	Lum>Geo
gleaming flat	0	Lum	Geo	Lum>Geo
radiant sleek	0	Lum	Geo	Lum>Geo
jet-black choppy	0	Lum	Geo	Lum>Geo
bright spiny	0	Lum	Geo	Lum>Geo
unilluminated circular	0	Lum	Geo	Lum>Geo
shady velvety	0	Lum	Geo	Lum>Geo
glowing slick	0	Lum	Geo	Lum>Geo
adorable spherical	1	Val	Geo	Val>Geo
lovely silky	1	Val	Geo	Val>Geo
alarming spiky	1	Val	Geo	Val>Geo
frightening glossy	1	Val	Geo	Val>Geo
spooky round	1	Val	Geo	Val>Geo
endearing rough	1	Val	Geo	Val>Geo
beautiful lumpy	1	Val	Geo	Val>Geo
pretty sharp	1	Val	Geo	Val>Geo
rough endearing	0	Val	Geo	Val>Geo
lumpy beautiful	0	Val	Geo	Val>Geo
silky lovely	0	Val	Geo	Val>Geo
spherical adorable	0	Val	Geo	Val>Geo
glossy frightening	0	Val	Geo	Val>Geo
sharp pretty	0	Val	Geo	Val>Geo
spiky alarming	0	Val	Geo	Val>Geo
round spooky	0	Val	Geo	Val>Geo

Note. Canonical pairs were determined by English adjective ordering rules and are indicated by a value of 1, while non-canonical pairs are shaded in red and correspond to a value of zero.