EFFECTS OF SPATIAL LANGUAGE CUES ON ATTENTION AND THE PERCEPTION OF AMBIGUOUS IMAGES

By

Aaron Foster

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Arts in Psychology: Academic Research

Committee Membership

Dr. Kauyumari Sanchez, Committee Chair

Dr. Amber Gaffney, Committee Member

Dr. Gregg Gold, Committee Member

Dr. Christopher Aberson, Program Graduate Coordinator

May 2021

Abstract

EFFECTS OF SPATIAL LANGUAGE CUES ON ATTENTION AND THE PERCEPTION OF AMBIGUOUS IMAGES

Aaron Foster

An ambiguous object, for the purposes of this thesis is any object that has more than one. interpretation to it. The brain is designed to "fill in the blanks" and make sense of the world. Thus, it will use anything available, like language, to help in resolving the ambiguity. Language can change how we perceive information in the world (Dils & Boroditsky, 2010) and where we direct our attention (Ostarek & Vigliocco, 2017; Estes et. al. 2008; Estes, Verges, Adelman, 2015). Language can play a role in resolving ambiguity by directing attention in certain directions. For example, if I say "upward." and you see something in the sky, you might be inclined to perceive items that are typical in that location (e.g., bird and plane) as compared to atypical items (e.g., wrench) (Estes, Verges, & Adelman, 2015; Estes, Verges, & Barsalou, 2008). However, to date, no study has investigated whether it is possible that such spatial language cues (like "upwards" and "downwards") can affect the interpretation of an ambiguous stimulus. The aim of this thesis is to explore the effect of spatial language cues on the perception of ambiguous images.

Acknowledgements

I would like to acknowledge my thesis advisor, Dr. Sanchez. Without her guidance I would not be where I am today. Dr. Sanchez's guidance, wisdom, and teaching got me to where I am today and is a part of the reason, I am the student I am today. Also, I would like to thank Dr. Maria Iturbide, who has been my advisor since I first joined Humboldt State as an undergraduate. Without the support and guidance of Dr. Iturbide, I would not have even joined the Academic Research program at Humboldt State. Even after joining the program, Dr. Iturbide still served as a beacon of support and guidance when called upon, making sure I did not stray from progress. A thank you to Dr. Gaffney, who has also supported me through my undergraduate career. Dr. Gaffney has been a source of encouragement and positivity throughout my time in the department, always keeping me positive and on track and offering as much help as she is able to give. My committee members were also a huge source of guidance and support on this journey. My friends, Bernardo, Doug, Lex, Lauren, Jake, Chelsea, and Stephanie have helped and assisted me with countless hours of edits, proofreading and practice presentations. The amount of time helped and assisted in coding the project, assistance in writing and editing kept me motivated to keep the project on track and finish the project on time. Without them, this thesis would not be half the document it currently is. A special thank you to my family for supporting me through this process, cheering me up when time(s) got dark and always pushing me to be the very best that I could be, even when I felt like I could not give any more to the program.

Table of Contents

Introduction

It's a bird! It's a plane! It's superman!? Sometimes there are things in our world that are ambiguous. An ambiguous object, for the purposes of this thesis, is any object that has more than one interpretation to it. The brain is designed to "fill in the blanks" and make sense of the world. Thus, it will use anything available, such as language, to resolve ambiguity. Language can change how we perceive information in the world (Dils & Boroditsky, 2010) and where we direct our attention (Estes et. al. 2008; Estes et al., 2015; Ostarek & Vigliocco, 2017). Language can play a role in resolving ambiguity by directing attention in certain directions. For example, if I say "upward" and you see something in the sky, you might be inclined to perceive unambiguous items that are typical in that location (e.g., bird and plane) as compared to atypical items (e.g., wrench) (Estes et al., 2008, Estes et al., 2015). However, to date, no research has investigated whether it is possible that such spatial language cues (e.g., "upward" and "downwards") can affect the interpretation of an ambiguous stimulus. The aim of this thesis was to explore the effect of spatial language cues on the perception of ambiguous images. 0

Perception, Attention, and Ambiguous Images

To date, only one study has investigated the role of language in perception and attention on ambiguous images (Dils & Boroditsky, 2010). Dils and Boroditsky (2010) found that attention and the perception of ambiguous stimulus can be influenced by movement cues. In three experiments, movement was manipulated: Via real visual motion in experiment one, via language that directly suggested motion in experiment two, and indirect (abstract) language in experiment three. In all experiments the ambiguous image came in the form of a single stimulus that when perceived upward it depicted a hawk, and when perceived downward it depicted a goose.

In experiment one, Dils and Boroditsky (2010) questioned whether real visual motion could affect attention and the perception of an ambiguous image. Participants were asked to watch a screen that depicted dynamic lines moving upward or downward. Immediately after, an ambiguous image was shown on the screen and, to check for perception of the ambiguous image, participants were asked to click on the beak of the bird they saw in the image. Although one could interpret two different birds (i.e., hawk or goose), these instructions attempted to avoid drawing attention to the ambiguity. Clicks above the midline were coded as an "upward" hawk interpretation and clicks below the midline were coded as a "downward" goose interpretation. After the participants indicated the beak location, they were asked if they had noticed anything more about the image that they had seen to exclude participants who noticed the ambiguity and the second interpretation.

The data from experiment one of Dils and Boroditsky (2010) suggests that participants who saw the dynamic lines moving upward perceived the image facing upward (the hawk interpretation) and the participants who saw the dynamic lines moving downward perceived the image facing downward (the goose interpretation). This finding suggests that looking at real visual motion could serve to direct attention to influence the perception of an ambiguous image.

With the second experiment, Dils and Boroditsky (2010) tested whether attention and the perception of an ambiguous image could be affected via language that directly suggested literal motion. Participants read a passage that described either upward motion (e.g., “the kids walked from the 9th floor to the 12th floor”) or downward motion (e.g., “the kids walked from the 20th floor to the 18th floor”), followed by a comprehension question related to the passage. After the story, participants were presented with the ambiguous hawk/goose image that was centered in the middle of the screen. Participants were instructed to click on the beak and draw a worm on the beak of the bird. Participants who drew the worm on the top of the image were coded as seeing the image from an “upward” hawk interpretation. Participants who drew it on the bottom were coded as a “bottom” goose interpretation. Overall, participants saw the image in the direction that was consistent with the story they read. For example, if participants read the “upward” story, they saw the image as a hawk over the goose. The opposite was the case for the “downward” story (Dils & Boroditsky, 2010). This suggested that literal motion could possibly serve to direct attention and that this could have affected the perception of an ambiguous image.

Dils and Boroditsky’s (2010) third experiment tested whether attention and the perception of an ambiguous image could be affected via language that indirectly suggested motion (i.e., abstract language). Participants read passages describing abstract upward motion (e.g., the stocks skyrocketed from $30 to $300) or abstract downward motion (e.g., the stocks plummet from $300 to $30). After reading, the participants would view the ambiguous hawk/goose image and draw a worm on the bird's beak, as in experiment two. In this case, the value of money is a separate concept from the words itself and was not the target of the experiment. The results showed that abstract language was not able to serve as a means to direct attention and did not affect the perception of an ambiguous image.

Summary of Perception, Attention, and Ambiguous Images

Overall, the findings of Dils and Boroditsky (2010) suggest that attention can be directed and that this can influence the interpretation of an ambiguous image. However, it seems that potentially only via real visual motion and language that directly suggested literal motion can achieve this shift in attention and perception. However, it was unknown whether words, such as spatial language cues, (e.g., "upward" or "downwards") were able to similarly affect the interpretation of an ambiguous image, as they were not explicitly used or tested for this effect, which was part of the purpose of this thesis.

Spatial Language Cues on Attention and Perception

Although Dils and Boroditsky (2010) provide some of the few empirical studies investigating the role of motion on attention and the perception of an ambiguous image, other studies have investigated the impact of spatial language cues on attention. For example, Estes et al., (2008) tested whether object words (e.g., hat) can serve as a spatial cue (e.g., directing attention upward). They also tested whether the direction of one's attention would affect the identification of a target object when the target is in the same place as one's directed attention (congruent), or in a different place as one's directed attention (incongruent).

In the first experiment, participants were first presented with a context cue (e.g., cowboy) to orient the spatial language cue (e.g., hat) that followed (Estes et al., 2008). The target object (i.e., "x") was shown after the spatial language cue at either the top or the bottom of the screen, both of which were equally likely. In half of the trials, the spatial language cue matched the location of the target object, while the other half of the trials there was a mismatch. Participants were asked to click on the target object as fast as possible. The results found that when the spatial language cue was congruent with the location of the target object, reaction times were slower compared to the mismatch trials. This means that when participants were shown "hat" and the target appeared at the top of the screen, response time was slower, and accuracy was less than those instances when participants were shown "hat" and the target appeared at the bottom of the screen. This suggested that interference occurs when one's attention is in the same location as a target object.

For experiment two, Estes et al. (2008) conducted a replication plus extension of the first experiment which simply removed the context word (e.g., cowboy). The data replicated the results from experiment one. When the spatial language cue (hat) was congruent with the location of the target object, reaction times were slower compared to the trials where there was a mismatch. This suggests that the interference was not caused by the context word, but the spatial language cue specifically. These results suggest that interference occurs when one's attention is in the same location as a target object. In addition, Estes et al. indicated that a single spatial language cue word is sufficient to shift attention.

The interference found in Estes et al. (2008) was surprising because of research that would predict results to the contrary, such as work on the spatial Stroop effect (Palef & Olson, 1975). In the spatial Stroop effect (Palef & Olson, 1975), facilitation and interference are found. For facilitation, when there is a match (congruent) between a spatial word (i.e., "above") and the location on the screen where the word is found (e.g., word is presented at the top of the screen) fast reaction times are observed. In contrast, when there is a mismatch (incongruent) between the spatial word (i.e., "above") and the location on the screen where the word is found (e.g., word is presented at the bottom of the screen) slow reaction times are observed comparatively. Thus, if the results of Estes et al. (2008) followed the Stroop like pattern, facilitation would be expected when the word "hat" is presented and the target "x" is at the top of the screen, but instead interference was found.

Estes et al. (2015) sought to investigate the interference effect observed by Estes et al. (2008) by using the perceptual matching theory. In the perceptual matching theory, when one is shown a noun that represents an object (e.g., hat), the brain then begins searching for the object (hat) in the typical area of space (i.e., upward, high in space) it is found. The brain also starts to mentally compare the object(s) found in the typical area of space to the noun. For example, if one is presented with the word 'hat' that person may be inclined to shift their attention upward in search of a hat, given that hats are typically found atop a person's head. If while searching, that person comes upon an object (i.e., a hat), they refer to the noun (i.e., hat) that started the search, creating the interference due to the object (i.e., the hat) not matching the object found (i.e., a hat). If the object found is related to the noun (e.g., if someone saw a hat and the noun they saw before was "hat") shown before and it matches the typical location of the object noun, facilitation occurs. If the object is not related to the noun (e.g., the "x" target in Estes et al. (2008)) and does not match the typical location, interference occurs.

To test perceptual matching, Estes et al. (2015), investigated the relationship between a spatial object cue (e.g., the word 'bird') and a target object that was a congruent real object (a picture of a bird), an incongruent real object (a picture of a wrench), or an incongruent abstract object (shapes and lines). Participants were first shown a blank screen, followed by a central fixation point, followed by a spatial object cue word (i.e., bird). After the spatial object cue word, a target object was shown at either the top or bottom of the screen, where it remained until participants indicated if the target was a real (i.e., bird or wrench) or abstract object (i.e., shapes and lines). It was found that conditions where the spatial object cue (bird) was congruent with the target object (i.e., picture of a bird) responses were faster and more accurate compared to conditions where the spatial language cue did not match the target object (wrench or abstract object), where instead interference was found. In addition, the data for the experiment show that when the spatial object cue word (i.e., "bird") was congruent with the target object (an image of a bird) at the top of the screen, reaction times were faster compared to conditions where the location (i.e., the bird image at the bottom of the screen) and the target object did not match (i.e., wrench or shapes and lines).

The results of Estes et al. (2015) suggest that the perceptual matching theory may account for the interference found in Estes et al. (2008). In Estes et al. (2008), participants were presented with a spatial language cue (i.e., hat) and were then asked to locate a target object (i.e., ‘x’). Given that the target object (i.e., ‘x’) is not related to the spatial language cue (i.e., hat), when the target object (x) was in the typical area of space (i.e., upward, high in space) it resulted in interference.

In a similar study, Zwaan and Yaxley (2003) tested whether words (e.g., “roof”) presented in typical (top of screen) or atypical (bottom of screen) referent positions would result in facilitation or interference (e.g., “Stroop like” effects). They found that pairs of words, like “branches”, when above “grass", their typical referent positions, were responded to faster than words that were simply related that don’t have typical referent positions (i.e., “house” above ``window"). In addition, they found that when words were in atypical positions, like “grass” above “branches”, that it resulted in the slowest response times. These results are in line with typical Stroop facilitation and interference effects. These results are also in-line with the perceptual matching theory, as words are responded to faster when in their typical spatial locations (branches above grass) compared to when in atypical locations (grass above branches). In atypical locations, the brain is searching for a match or something related to the targets, and when the brain fails to find a match or encounters a mismatch, this results in extra processing time to work through the incongruence.

Additional research has been done to further explain the results found in Estes et al. (2015). Some research shows that the interference from Estes et al. (2015) could be due to priming. Ostarek and Vigliocco (2017) questioned whether the interference effect found in Estes et al. (2015), where participants took longer to identify a target object (e.g., X) after being shown an object word (e.g., boot, hat), is driven by priming. They hypothesized that by reading the object word (e.g., boot, hat), the brain begins to simulate the actual perception of the object (e.g., hat, boot) in visual brain areas (Ostarek & Vigliocco, 2017). Because of this, when someone comes upon an object that is not related to the object word, interference occurs, and when someone comes across an object that is related to the object word, facilitation occurs. The main prediction was that prime-target pairs belonging to the same event (e.g., bird and cloud), should facilitate faster reaction time compared to conditions where they are not similar (e.g., bird and ground) which should inhibit reaction time. While three experiments were conducted, only the first will be described due to its relevance to the current study.

In the first experiment, Ostarek and Vigliocco (2017) asked participants to attend to prime words (e.g., bird, cloud) and were told they should be remembered for a memory test. After the prime word, a picture would appear either at the top of the screen or bottom of the screen and participants were told to indicate what the picture was as fast and accurately as possible by pressing a button. Results of experiment one show that seeing the prime word (e.g., cloud) and seeing the picture of a cloud at the top of the screen resulted in faster response time compared to seeing the prime word (e.g., cloud) and seeing the cloud at the bottom of the screen. Researchers propose that when reading the prime word (i.e., cloud), the brain and body reacts in an automatic way and the eyes naturally want to look up as if looking for a cloud in the sky, because that is where one's attention has been drawn, thus leading to facilitation when the target is located. These results are in line with the perceptual matching theory. When shown an object word (e.g., hat) the brain begins to scan for that object (e.g., hat) in its typical area (e.g., upward) until it finds the object, or comes upon something within that area (e.g., bird, cloud) that matches or is associated with the object word (e.g., hat) or does not match the object word (e.g., ground). The results could assist in explaining potential results found in Estes et al. (2015). However, such effects were avoided in the current study as object words were not used to prime objects and instead spatial language cues (i.e., upward and downward) were used to direct attention to an area in space.

Beyond the role of directing attention to one area or another, language can also affect the perception of objects in different orientations (Stanfield & Zwaan, 2001). Stanfield and Zwaan (2001) investigated whether reading a sentence that implies orientation of an object, the implied orientation would be part of the mental representation of the object. For example, they tested whether the perception of an object (e.g., upright hammer) is affected when the orientation of the object is made salient in a sentence prior to seeing the object (i.e., read about horizontal: hammer lying down or vertical: hammer upright). Participants were presented with a sentence (i.e., the nail was facing upward) depicting a physical everyday object (e.g., hammer, pencil) in a specific orientation (e.g., horizontal: hammer lying down, or vertical: hammer upright). The participants were then shown a fixation point, followed by a picture of a physical everyday object in a specific orientation (e.g., horizontal or vertical). The participants were told to indicate if the picture was congruent with the object in the previous sentence, indicating "yes" or "no" as quickly as possible. The data show that recognition of an object was influenced by orientation. Participants were faster and more accurate in recognizing a picture of an upward hammer after reading a sentence implying an upward hammer, compared to when reading a sentence implying a horizontal laying hammer. More specifically, when the mentioned orientation was congruent with the orientation of the picture shown, response time was faster compared to instances where the picture did not match the orientation (Stanfield & Zwaan, 2001). As a follow-up to this study, Stanfield and Zwaan (2002) tested to see if object orientation was a part of the mental representation that we create when we think about objects. The results show that recognition of an object is faster when the orientation described matches the object. This could also be the case with an ambiguous object and a spatial location cue. If the spatial location cue (e.g., "upward") and the rotation of the ambiguous stimulus matched (i.e., a hawk is shown facing upward and the goose is facing downward), then the spatial location cue (e.g., "upward") could influence how participants saw the ambiguous object (i.e., seeing the object facing upward).

As a follow-up to Stanfield and Zwaan (2001), Stanfield and Zwaan (2002) questioned whether an object that was mentally represented was represented with or without orientation. That is, when people mentally create objects, is orientation of that object a factor when we view objects in the world or is the orientation static. Stanfield and Zwaan (2002), conducted an experiment to test, and extend the literature on two competing theories on this question, the perceptual symbols theory, which is different from perceptual matching in that perceptual symbols theory relates to the mental representation of objects and their orientation, whereas perceptual matching does not. The perceptual symbols theory suggests that people activate and manipulate symbols during the comprehension of language, such that when orientation is implied in a sentence of the object, the mental representation of that object is a part of that mental representation. For example, a sentence such as “a hammer is upside down” when read would have the listener mentally represent that object as upside down, not right-side up. The other theory is the amodal theory, which states that when someone mentally represents an object, they do not mentally create the object with orientation in mind, they simply mentally create the object and ignore the orientation all together. Thus, orientation would not be present in the mental representation of the object and should not be affected when you see the object in the outside world. Because of this, according to the amodal theory, only facilitation would occur when presented with a sentence such as “a hammer is upside down”, regardless of the orientation of the picture of the hammer. Thus, a mismatch effect would not be expected.

In the first experiment, Stanfield and Zwaan (2002) sought to first demonstrate evidence for the perceptual symbols theory before comparing it to the amodal theory. In this experiment, participants saw a sentence (e.g., she cooked the egg in the pan) on a screen that either did or did not indicate an object they would see (i.e., an egg). After reading the sentence, participants were asked to indicate if they understood the sentence by pressing the spacebar on the keyboard in front of them, followed by a brief fixation point, followed by a picture of the target object (i.e., egg). It was found that participants were faster to respond to the picture of the target object (i.e., an egg) if the sentence that indicated the object (i.e., "She cooked the egg") matched with the picture shown (e.g., a cooked egg vs an egg in a shell). In contrast, participants were slower to respond to the picture of the target object if the object was not mentioned in the sentence compared to the condition where the image (e.g., an egg) was mentioned in the sentence.

According to Stanfield and Zwaan (2002), the results of the data support perceptual symbols theory. They theorize that when reading about objects in sentences, participants also represented the implied shape of the object as well, so when the shape did not match, response time was slower (Stanfield & Zwaan, 2002). Interference occurs when the opposite was true. When the implied orientation in the sentence did not match the object shown later, response time was slower due to the incongruence between the two, as per perceptual reference theory.

In the second experiment, Stanfield and Zwaan (2002) investigated further for support of either the perceptual symbols theory or amodal theory by simply naming the picture they saw after reading a sentence. The procedure for the experiment was similar to experiment one, except participants named the object (e.g., "hammer" to the sentence "a hammer is upside down") instead of indicating if the object appeared in the sentence. The results found that when a sentence referenced an object (i.e., a hammer is upside down") and participants were then shown a picture of an object (i.e., hammer), that reaction times were faster when the orientation of the object in the sentence matched the orientation of the object in the picture. These results support perceptual symbols theory. Given the slower reaction time of the participants when presented a mismatched case between sentence and pictured objects, researchers conclude that we represent orientation of objects when we mentally create them. This finding supports the perceptual symbols theory, due to mentally representing orientation and causing a mismatch effect. When the conditions were not congruent (e.g., orientation did not match the picture), reaction time was affected, thus showing a mismatch effect. These results do not support the amodal theory, as the results show that when orientation matches, reaction time is faster compared to when orientation does not match. Taking this notion one step further, it might be possible to sway people into mentally representing an object in a certain orientation by using words like 'upward' and 'downwards'.

While previous research has shown the effect of object words and their effect on attention, the use of simpler words needs to be further emphasized, given the relevance to the current project. Hommel et al. (2001) tested to see if over-learned communicative signals, such as an arrow pointing upward or downwards or words such as “up” and “down,” were an effective way to direct participants' attention to a specific area.

Hommel et al. (2001) theorized that congruent conditions, when the stimulus location was congruent with the cue (i.e., when the arrow predicted the location of the target), reaction time would be faster compared to when the arrow did not predict the location of the target (i.e., “x”). The trial(s) began with a central fixation point and four grey boxes in one of the four locations. The fixation point was then replaced by one of two spatial location cues, either an arrow pointing in one of the four directions (e.g., ->), or a word that also indicated one of the four directions (e.g., right). Participants searched for a target letter (e.g., “x”) in one of four locations (left, right, up, or down) and were to indicate which box it was located in. The results show that reaction time was faster when the target location and spatial location cues were congruent in direction, demonstrating typical Stroop facilitation effects. Results also show that when the target location and the spatial location cues were incongruent, participants were slower, or prone to more errors when searching for the target. The data suggest that symbols, such as arrows can be used to direct attention to specific areas on screen. While the effects of object words and simple symbols have been identified, it is also worth noting that verbs and other types of words have also been shown to have the same effect.

Summary of Spatial Language Cues on Attention and Perception

In this section, results have shown that object words, spatial location cues and orientation have an effect on locating the target object shown, as well as affecting reaction time, and location as to where participants look for the target (Ostarek & Vigliocco, 2017; Dils & Boroditsky, 2010; Estes et al., 2015; 2008; Hommel et al., 2001; Stanfield & Zwaan, 2002; 2001;). Also, results have shown that, potentially, the spatial Stroop effect can also affect the outcome of a location search, given whether the location cue and the object are congruent or not (Palef & Olson, 1975). The spatial Stroop effect also shows an effect on the reaction time of participants locating the target. In many conditions, if the location cue and the target are congruent with each other, participants tend to be faster in locating the target compared to when the location cue and the target are incongruent with each other, showing slower reaction times (Palef & Olson, 1975). Lastly, research also shows that orientation can influence reaction time when attempting to locate the object (Stanfield & Zwaan, 2001; 2002). When orientation is implied within a sentence, participants are faster to locate the object when it matches the orientation that participants read about. The opposite effect is found when the orientation does not match the target object. Reaction time is slower when the sentence that implies orientation and the target object orientation do not match. The results that are most consistent with the current study are the spatial Stroop-like effects shown in Palef and Olson (1975) and are used to support the hypotheses used in the current study.

Spatial Language Cues on Perception: Embodied Cognition Approach

The impact of spatial language cues on perception may potentially be explained by the theory of embodied cognition (Glenberg & Kaschak, 2002). Embodied cognition, according to Wilson and Glonka (2013) is the theory that our brains are not the only resource responsible for our ability to respond to the environment. According to Wilson and Glonka (2013), our bodies, the availability of resources, and the motion available to our body is used to interact with the environment. Wilson and Glonka (2013) also propose that language plays a role in embodied cognition.

According to Glenberg and Kaschak (2002), an example of embodied cognition can be found within words that imply action (e.g., stand, sit, breathe) and that the meaning of these words is a part of the words themselves (e.g., embodied), and the only requirement for understanding the words is that the words be placed in a sentence is such a way that the words make sense to the person reading it. For example, using the spatial language cue “upward,” one could theorize that the word implies looking up, or shifting attention in the upward direction. As long as the word “upward” is placed in a sentence that makes sense (e.g., the kite rose upward), Glenberg and Kaschak (2002) suggest that the meaning of the action word (e.g., upward) makes sense in context. In fact, the results of Dils and Boroditsky (2010) can be accounted for using an embodied cognition approach. For example, in the experiment by Dils and Boroditsky (2010), participants were affected by the simulated movement implied by the story they read and the story made sense from start to finish. In the story, an alternative explanation presents itself. The simulated movement was achieved by imagining one was traveling in an elevator going from the 19th to the 20th floor (implying upward movement) or 20th floor to 19th floor (implying downward movement). This resulted in participants identifying the image facing upward, if given the story of the elevator traveling upward, and identifying the image facing downward, if the participants were given the story of the elevator traveling downward.

A subset of the embodied approach is the indexical hypothesis. The indexical hypothesis contends that meaning is based on action (Glenberg & Kaschak, 2002). The meaning of any situation is made up of a set of actions available to someone at a given time (Glenberg & Kaschak, 2002). For example, if you are in a room with a chair, the affordances of the chair may change depending on the situation. Affordances are potential actions between body (yourself) and the object in question (e.g., chair). Wilson and Glonka (2013) contest that actions (e.g., moving, grabbing, reaching) are a core part of the embodied approach. For instance, when a tennis player intercepts the tennis ball, Wilson and Glonka (2013) contest that instead of some elaborative inner formula being conducted, the solution to the problem (e.g., getting from A to B) is as simple as an action (e.g., running toward the ball and hitting it) and the object in question (e.g., the tennis ball). If the tennis player cannot complete the action (e.g., hitting the ball), then other available options are used instead (e.g., giving up the ball). Thus, if your legs were feeling tired, the chair may afford sitting. However, if someone wanted to reach for something really high up, the chair would afford someone a higher place to stand, to serve their needs in the given situation. And if they were in an argument, the chair could afford throwing.

The indexical hypothesis implies that the meaning or action implied by action words (e.g., sit and stand) are already embodied or a part of the word itself. According to the indexical hypothesis, as long as the action word makes sense in context (e.g., You [gave] the cards to Susan) then the meaning of the action word (e.g., give / gave) is known as long as the sentence makes sense in context. If the word does not make sense (e.g., You [sang] the cards to Susan), then the meaning of the action verb (e.g., sang) is not known, because it does not make sense within the context of the sentence (Glenberg & Kaschak, 2002). The indexical hypothesis is attempting to counter the idea that to understand the meaning of a word, anyone is required to learn each specific symbol (i.e., the letters) of the word first and bring them together before meaning can be "mapped" to that word (e.g., understanding that "d," "i," and "g" together make the word dig; Glenberg & Kaschak, 2002).

According to Glenberg and Kaschak (2002), there are three processes that turn words and syntax into an action-based meaning. The first is the theory that words and phrases are mapped to perceptual symbols. Perceptual symbols are based on words and phrases that are mapped to them (e.g., a physical chair and the word "chair"). Secondly, affordance, or the meaning of words are derived from perceptual symbols (e.g., words) and are flexible in nature and thus words (e.g., object words) and the use of those words are not necessarily related to the object they imply. Affordances are potential actions between body (yourself) and the object in question (e.g., chair). For example, "hang your hat on the upright broom" can be judged as making sense, as we can see that the word "broom," while it is used to sweep, could also be used as a coat rack if it is upright. In this case, the affordance (action between body and object) would be the "hang." Lastly, the indexical hypothesis states that affordances (e.g., different meanings) are brought together under syntactic constructions, or when the sentence makes sense. For example, "please hang your coat on that cup" would be coded as not making sense, thus the meaning of the action verb itself (e.g., hanging a coat on a cup) is not understood, thus the sentence is nonsensical. Only when a sentence makes sense with the action verb (e.g., hanging) does understanding take place (i.e., please hang your hat on the vacuum / cup). In the example where the sentence makes sense, the meaning of the action is understood, and the action can take place without interference.

Glenberg and Kaschak (2002) questioned whether the meaning of a word (e.g., push) was understood through having to learn each individual character (e.g., "p","u","s", and "h") or if the meaning of the word was embodied in the word itself, as long as the word made sense within context. They hypothesized that when the direction of movement in a sentence matched the same direction as the movement produced by the participant, that facilitation (quicker movements) would occur, while a mismatch between the sentence and real movement would lead to errors and slower movements. For example, according to the indexical hypothesis, sentences that are sensible and imply direction away from the body (e.g., I push the drawer closed) should interfere with contrary movements, such as when participants are asked to pull a button close to them. In contrast, sentences that are sensible and imply direction close to the body (e.g., I pulled the drawer open) should lead to facilitation when participants are asked to pull a button close to them. In their study, Glenberg and Kaschak (2002) gave participants either a sentence that made sense (e.g., open the drawer) or nonsense sentence (e.g., change the heat from the cup). The sensible sentences depicted an action in a certain direction, either toward the body (e.g., open the cupboard, put your hand on your stomach) or away from the body (e.g., close the cupboard, put your hand out in front of you). The goal of the sentence was to imply movement in a direction. After which, participants indicated if the sentence made sense by pressing one of two buttons, one button was close to their person and one button was far from their person. A third button was neutral, and participants started on this button to display the sentence. The participants were never instructed to pay attention to the implied direction. In the first experiment, participants were instructed to move their hand from the middle button to the farthest button if the sentence made sense (e.g., put your hand on the cup), this would imply movement away from the person), and to the closest button if the sentence did not make sense (e.g., put your hat on the air on top of you), this would imply direction toward the person. The sentences could imply either direction moving toward or away from the person. In the other condition, participants were asked to move their hand to the button closest to them if the sentence made sense or push the button farthest away if the sentence did not make sense (e.g., put your hand on air).

According to Glenberg and Kaschak (2002), results indicate that the interaction between response direction and sentence direction was significant. When the sentence direction and response direction matched, it resulted in faster reaction times (RTs) compared to when there was a mismatch between sentence direction and response direction. These results support the indexical hypothesis because when participants were given a button that was farther away from their person and the sentence implied direction away from the body (e.g., put your hand on the cup), without being told to pay attention to the direction in the sentence, participants were inherently faster compared to the condition where participants were given a sentence that indicated movement away from the body (e.g., put your hand on the cup) and the button was closer to their person. According to Glenberg and Kaschak (2002), this shows that some language-understanding taps into the system responsible for actions.

Experiments 2a and 2b by Glenberg and Kaschak (2002) replicated and extended the first experiment by changing the hand the participants answered with from right hand to left hand (2a), and then went further to instead only a finger was used to give an answer, not the entire hand (2b). Both experiments 2a and 2b demonstrated similar results to those found in the first experiment; response time was faster when movement was congruent (e.g., away from the person) as implied in the sentence (e.g., put your hand on the cup) was not only just for the dominant hand (experiment 2a) and that the effect is not based solely on the spatial location of the responses possible, but the action it implies as well. With this, we see some support that language may not control our actions overall, but it may have some influences in our actions, depending on what we read before we act. The results support other potential reasons why spatial language cues might change our attention, based on what the spatial language cue is implying (i.e., upward vs downward). The indexical hypothesis provides support for explaining why some words, such as spatial language cues, have the effect of changing attention. According to the indexical hypothesis, language is embodied in the word and as long as that word (or sentence) makes sense, action is based on the word. While “upward” and “downward” were single words without context, they still make sense as sensible words, thus the direction that the word implies should hypothetically cause an action to occur based on the implied direction. These results suggest that, potentially, even without bringing attention to the words themselves, that spatial language cues could influence attention.

While we have seen evidence of attention being affected by spatial language cues, perception is also affected by spatial language cues. For example, Meteyard et al. (2007) showed that words affect perception, using a random dot kinematograms (RDK), a machine that generates dots and moves them in a random fashion, and reaction times as a type of measure. Participants were placed in a room where they listened to motion verbs (i.e., lift or drop). Participants were then shown an RDK that were either moving in the direction indicated by the word or moving randomly. The participants were told to ignore the words they listened to and indicate if the dots were moving coherently. It was found that the motion verbs (i.e., lift or drop) that the participants heard, even when told to ignore them, affected participants' perception of the dots moving. For example, those who heard a verb indicating upward motion (lift) saw the dots as moving upward in a coherent manner and the participants who heard a verb indicating downward motion (drop) saw the dots moving downward in a coherent manner. The direction of the dots moving coherently or randomly was counterbalanced, all presented in a random order. This study provides support for the use of spatial language cues and their impact on perception. This study, while not explicitly testing for the embodied approach, shows that words that imply meaning (e.g., drop, lift), still have an effect on the participant. One could argue that because the word makes sense by itself (e.g., drop), the actions available (e.g., to look down, perceive downward movement) are still available to the participants, even implicitly. Thus, the participants see movement that is coherent with the movement implied by the word.

Summary of Spatial Language Cues on Perception: Embodied Cognition Approach

Studies have shown that a single action word (e.g., drop) or spatial language cue (e.g., upward) are enough to have an effect on someone's perception of moving objects (Meteyard et al., 2007) or our own body movements (Glenberg & Kaschak, 2002, Wilson and Glonka, 2013). Meteyard et al. (2007) provides support for the theory that spatial language cues (e.g., drop, lift) on their own are enough to affect perception of an object moving coherently or incoherently. Glenberg and Kaschak (2002) and Wilson and Glonka (2013) attempt to provide a theory for these types of results by using the theory of embodied cognition, and within it the indexical hypothesis. The theory being that action is based in meaning, and that as long as the word is understood in context (e.g., the rain falls from the sky), it is unnecessary to understand each character independently (e.g., "r", "a", "i", "n") and the meaning of the sentence is already known (Glenberg & Kaschak, 2002; Wilson & Glonka, 2013). In the indexical hypothesis, action (e.g., sit) toward an object (e.g., chair) is based on the availability of the action (e.g., sitting) at the given time. Also, according to embodied cognition, the indexical hypothesis that the meaning of the action (e.g., sit) is already understood (e.g., sitting in a chair), as long as the object (e.g., chair) is connected to or makes sense with the action.

Attention on the Perception of Ambiguous stimulus

Attention and ambiguity separately have shown results relevant to the current study. Bernstein and Cooper (1997) questioned whether interpretation of an ambiguous stimulus would be influenced by direction of motion of the stimulus. The hypothesis was that direction of motion (up, down, right and left) would influence interpretation of an ambiguous stimulus. Participants viewed an ambiguous stimulus (rabbit/duck) whose interpretation is changed from one to the other horizontally (right/left orientation). This ambiguous stimulus (e.g., rabbit/duck) was presented within the context of a frame (e.g., rectangle). The frame would move position, either leftward or rightward, in order to simulate motion of the object within it. Participants would then indicate the interpretation (rabbit/duck) of the ambiguous image they saw. Results of the experiment show that interpretation of the ambiguous stimulus was influenced by the direction of motion. Participants who saw the frame moving rightward interpreted the image facing right (rabbit). Participants who saw the frame moving leftward interpreted the image facing left (duck). The results of the first experiment show support for real motion in influencing perception. However, it was unknown whether spatial language cues could serve to influence the perception of an ambiguous image.

Bernstein and Cooper (1997) have shown that the frame that a stimuli is placed in can have an effect on what is perceived if a stimulus is placed in different settings. It is also worth noting that the frame does not need to move for the frame to have an effect on perception of an ambiguous stimulus. For example, Chen and Scholl (2014) investigated the role of the inward bias effect in three experiments. The inward bias effect is the tendency to view items as facing inward (toward the center) instead of outward when placed in a frame. In the first experiment, participants saw an ambiguous object (e.g., a rabbit/duck image) in the inner portion of a rectangular frame to see if participants would still interpret the object as facing inward. Participants were shown an ambiguous stimulus (e.g., a rabbit/duck image) toward the left or right side of a rectangular border/frame. They indicated if the image was either a duck or a rabbit, not being informed of the ambiguous nature of the image. Researchers suggest that information about the potential direction of motion is a source of information used to help us perceive or work through ambiguity. Bernstein and Cooper (1997) conclude that direction of motion is just one of many other sources of information that is used in order to help us understand what we are looking at or perceiving in our world.

It could also be that the actual eye gaze, or where one chooses to look at an image at the beginning determines what one sees in an ambiguous image. There is another line of research that suggests that eye gaze may also influence what you see in an ambiguous image, depending on where you look at the image. Einhäuser et al. (2004) questioned whether eye gaze and eye movement were directly related to seeing different perceptions in an ambiguous stimulus, specifically in a Necker cube (e.g., an ambiguous cube with negative space). The hypothesis was that perceptions (e.g., perceiving the Necker cube facing in a certain manner) of the Necker cube, would change (e.g., seeing different depth cues, or perceiving different parts of the object as facing forward) based on eye movement (e.g., where you were looking at the cube) and eye gaze. Einhäuser et al. (2004) instructed participants to view a Necker cube and report their percept of the Necker cube by pressing one of two buttons on a mouse. Participants' eye gaze was recorded and each trial began with a central fixation point in the middle of the screen to form a baseline. Apart from the fixation point, no other item influenced the movement of the participants' eyes nor was eye gaze mentioned to the participants before the experiment. Experimenters recorded changes in percepts that were indicated by participants and matched them with eye gaze recorded at the time the participants indicated a change in perception of the Necker cube.

According to Einhäuser et al. (2004), results indicated that the changes in perception of the Necker cube occurred concurrently with changes in eye movements. Meaning that when participants indicated that a change in perception of the Necker cube occurred, eye movements indicated a change in eye gaze on the Necker cube. Researchers indicate that this experiment only shows correlational results, and that causation should not be implied by these results. Results show that, potentially, depending on where participants choose to look at an ambiguous stimulus first, their interpretation of the ambiguous stimulus could be altered, ignoring all other effects such as spatial location cues. However, it could also be that based on where the eyes land first, based on the spatial language cue, perception could be biased to that location and thus the interpretation could be different.

Current Study

The current study brought perception, attention, and language together to assess their effects on ambiguous images. Dils and Boroditsky (2010) showed that language via passages that depict movement can influence the perception of an ambiguous image. However, they did not use simple spatial language cues like ‘upward’ and ‘downwards’, nor did they vary the location of the images, as they were only presented in the center of the screen.

In addition, the research of Estes et al. (2008) suggested that spatial language cues could serve to shift attention with words like “upward” and “downwards” but had not been investigated with respect to the perception of ambiguous images. Moreover, the work by Bernstein and Cooper (1997) and Einhäuser et al. (2004) suggested that the movement of one’s eyes could affect the perception of an ambiguous image but had not been investigated with respect to spatial language cues. The current study sought to investigate the effect of spatial language cues on attention and the perception of ambiguous images.

Hypothesis/Predictions

Hypothesis 1a: Image location X Spatial Language Cue: Ambiguous Stimulus: Interpretation

If image location and verbal spatial language cues affect ambiguous stimulus interpretation, then differences in ambiguous stimulus interpretation should be observed when changes in the levels of image location and cue word occur. Specifically, when the image location is on the top of the screen and the spatial location cue is upward and thereby congruent participants should interpret the ambiguous stimulus as more “hawk-like” than “goose-like.”

Similarly, when the image location is on the bottom of the screen and the location cue is downward and thereby congruent, participants should interpret the ambiguous stimulus as more “goose-like” more than they report seeing “hawk-like.”

However, when the conditions are incongruent, in that image location and location cue do not match up in any condition, ambiguous stimulus interpretation will be equal and both precepts of the image will appear equally.

This hypothesis is supported by the results found by Dils and Boroditsky (2010) who found that participants who read passages implying upward movement were more likely to see the image facing upward compared to the image facing downward, and those who read passages indicating downward movement were more likely to see the image facing downward compared to the image facing upward. The hypothesis is also supported by the results found in Estes et al. (2008) and Hommel et al., (2001) which shows that spatial language cues, or object words, can influence the direction participants tend to look first.

Hypothesis 1b: Main Effect: Image Location: Ambiguous stimulus: Interpretation

If ambiguous stimulus interpretation is influenced by differences in image location, then differences in ambiguous stimulus interpretation should be observed when changes in the levels of image location occur. Specifically, when the image location is on the top of the screen, participants will report seeing a hawk facing upward compared to when the image is on the bottom of the screen where they would report seeing a goose facing downward. This is in line with Dils and Boroditsky, (2010) and is also supported by Stanfield and Zwaan (2001) which showed that participants were faster and more likely to interpret the object that matched the orientation prime shown before the item or image.

Hypothesis 1c: Main Effect: Spatial Language Cue: Ambiguous stimulus: Interpretation

If ambiguous stimulus interpretation is influenced by differences in spatial language cues, then those with the "UPWARD" condition of location cue will report seeing the hawk in the image compared to those in the "DOWNWARD" condition of location cue who will report seeing the goose more. This is in line with the results found in Dils and Boroditsky (2010) who found that when participants were given the word "up," they reported the image facing upward more so than the image facing downward. The opposite is true when participants were given the word "down." When participants read “down,” participants reported the image facing downward more so than the image facing upward.

Hypothesis 2a: Interaction: Image location X Spatial Language Cue: RT

If image location and spatial language cues affect reaction time, then differences in reaction time should be observed when changes in the levels of image location and cue word occur. Specifically, when the image location is on the top of the screen and the location cue is upward and thereby congruent participants should be faster to respond then those in the downward condition and thereby incongruent.

Similarly, when the image location is on the bottom of the screen and the location cue is downward and thereby congruent, participants should also be faster in the congruent downward condition than the incongruent downward condition.

However, When the conditions are incongruent, in that image location and location cue do not match up in any condition, response time will be slower for both image location and location cue. This hypothesis is in-line with the results found in various other studies (Palef & Olson, 1975; Hommel et al., 2001; Estes et al., 2015) who found spatial Stroop effects within their results. If the target object (e.g., X) was in a congruent location (e.g., top of the screen) with the spatial language cue (e.g., upward), then reaction time was faster compared to the target object (e.g., x) being in an incongruent location (e.g., bottom of the screen) with the spatial language cue (e.g., upward), showing slower reaction times when this was the case. Hommel et al. (2001) show similar results using spatial language cues in the form of arrows. Their results show that when the arrow pointed toward the target object (e.g., x), reaction time was faster compared to when the arrow pointed away from the target (e.g., x) in which reaction time was slower.

Method

Participants

IRB approval (IRB 20-036) was obtained on November 18th, 2020. Based on a power analysis using G-Power, the study required 133 participants to reach power of .30. However, only 57 participants (61% female, 21% male, 12% non-binary, 5% prefer not to answer) were recruited at Humboldt State University via the SONA subject pool, Psychology courses at HSU, Psi Chi's online recruitment portal, and snowball sampling methods. Participant ages were between 18-35+ years-old. Age was coded as a nominal variable, thus 52% were between the ages of 18-21, 25% between 22-25, 8% between 26-29, 10% were 35+, 4% preferred not to answer. In terms of racial demographics, participants were 50% White, 17% Latinx, 12% Multi-Racial, 10% Asian and 1% African American. 5% of participants chose not to answer. Participants received Psychology courses extra credit for participating in this study. All participants had normal or corrected to normal vision.

Materials

Ambiguous image. One ambiguous stimulus (Dils & Boroditsky, 2010) was used for this experiment. This image was manipulated using a vector drawing technique. This was done so the image would maintain its resolution quality on any computer screen (e.g., different ratios, resolutions) it was presented on. See appendix A.

Spatial language cues. Two words, 'UPWARD' and 'DOWNWARD' served as spatial language cues, from Estes, Verges, and Barsalou (2008).

Demographics. For the purposes of describing the sample of college students collected from, demographic information such as gender identity, age, and ethnicity was collected.

Design

The study was a 2 (Image location: top of the screen or bottom of the screen) x 2 (Spatial language cue: ‘UPWARD’ or ‘DOWNWARD’) between subjects design. The Image location cue variable was a between-subjects with two levels: top and bottom. In the top condition, the target image was at the top of the screen, while in the bottom condition, the target image was at the bottom of the screen, The Spatial language cue variable was also between-groups with two levels: upward and downward. The upward condition was noted by use of the word “UPWARD”, while the downward condition was noted by the use of the word “DOWNWARD”, prior to viewing the target image.

The dependent variables in this study were ambiguous stimulus interpretation and reaction time. For ambiguous stimulus interpretation, participants were instructed to click on what they believe to be the mouth of the bird they saw in the picture. This was used as an indicator for the participant’s interpretation of the ambiguous stimulus. The reaction time measure was measured from the onset of the target image to the point the mouse was clicked to indicate the location of the beak.

Procedure

The experiment was conducted via a website called Cognition.run (Cognition), which hosted the experiment. Participants engaged in the experiment from their own personal computer upon accessing the experimental website link. Upon acquiring their informed consent, participants were presented with the experiment instructions on their computer screen. The instructions informed the participant that they were being tested for the perception of bird silhouettes and that they should click on the beak of the bird's mouth as quickly as possible. Participants were randomly assigned to conditions.

The experiment consisted of a single trial. At the start of the trial, a spatial location cue word ('UPWARD' or 'DOWNWARD') was presented on the screen for 500ms. Immediately after, an ambiguous image appeared either at the top or bottom of the screen and stayed on the screen until the participant clicked on the animal's mouth. The image locations and spatial language cues were randomly placed (i.e., not counterbalanced) between participants. Participants were thanked for their response and were asked to answer demographic questions before being debriefed.

Table 1

Logistic Regression Results

Variables	*z*	*df*	*p*	*LR*	*p*	*b*	*Exp(b)*
Image Location	1.04	2	.05	6.0	.05	-1.9	3.6
Word Condition	-1.74	2	.18	3.5	.18	1.28	.15
Word Condition X Image Location	0.13	3	.90	7.46	.06	0.20	1.22

Results

Data Analysis

All data was collected using the website hosting the experiment, Cognition.run (Cognition) and the data were processed using R (R Core Team, 2021). The dependent variables in this study were ambiguous stimulus interpretation (hawk or goose) and reaction time.

For the final analysis a logistic regression, using the package “Stats” (R Core Team, 2021) in Rstudio (Rstudio Team, 2020) was run against the data in order to test the categorical dependent variable (e.g., ambiguous stimulus interpretation), while a regression model, using the package “Stats” (R Core Team, 2021) in Rstudio (Rstudio Team, 2020) was run to test the reaction time variable, which was the other dependent variable. Due to only having a single data point, the dependent variable being categorical, and the investigation of main effects and interactions, a logistic regression model fits the data best as it is structured to test single-data-point experiments and compare them for significance. The reason why a Chi-Square analysis was not used to analyze the data was because a Chi-Square was limited in the provided interpretation. A Chi-Square could not inform us of main effects between the independent variables (e.g., image location; spatial location cue) and the dependent variable (ambiguous stimulus interpretation).

Data Scoring

For the purposes of the experiment, the image used was a vector drawn image and presented in a singular rotation (see Appendix A). The image was coded so that only a click on the heads of the image (e.g., the hawk or the goose) would advance the experiment and be coded. This was done to avoid people clicking on certain areas of the image to indicate what they say and for specificity of data collection. If participants clicked on the top of the image (e.g., the hawk), their response was coded as “top.” If participants clicked on the bottom of the e.g., (e.g., the goose) their response was coded as “bottom.”

Exclusionary criteria

For the purposes of exclusion, participants were asked if they had seen the images before. Participants who had responded yes were excluded from final data analysis. This resulted in five data points being dropped from the final analysis for this reason. Also, participants who took more than 2,000ms. (Dils and Boroditsky, 2010) were excluded from the final analysis. This resulted in 24 participants being dropped from the final analysis due to exceeding 2,000ms reaction time Due to a programming “bot” obtaining the link, all records of the experiment that indicated a “bot” had attempted to access the link were deleted and not used in the final analysis, resulting in the removal of 212 “participants". After exclusionary criteria were implemented, the data reflected 48 participants: 13 participants in the Image location: Bottom of the screen, Spatial language cue: DOWNWARD condition, 13 participants in the Image location: Bottom of the screen; Spatial language cue: UPWARD condition, 6 participants in the Image location: Top of the screen; Spatial language cue: DOWNWARD condition, and 16 participants in theImage location: Top of the screen, Spatial language cue: UPWARD condition.

Preliminary Analysis

Before the primary analysis, kurtosis and skewness was investigated against the reaction time variable using the DescTools package (Signorell, 2021) in Rstudio (Rstudio Team, 2020). Additionally, further assumptions of variance were assessed as well. Reaction time was assessed using the Stats and Graphics (R Core Team, 2021) package in Rstudio. Visual inspection of the graphs indicated that reaction time was normally distributed and not violating any regression assumptions. Due to the independent variables being qualitative, no correlation tests were examined on the variables themselves.

Hypothesis 1a: Image location X Spatial Language Cue: Ambiguous Stimulus: Interpretation

It was expected that there would be an interaction effect between image location and spatial language cues on the interpretation of the ambiguous stimulus. After conducting a logistic regression, results of the model, which can be found in Table 1, indicate that the interaction between image location and spatial location cue was not significant. A chi-square test was run against the model to test for significance. The results indicated the model was not significant χ^2 (3; N = 40) = 7.46, p =.06 and an effect size (odds ratio) of 1.22.

Results indicate that no matter where the image was located on the screen (e.g., top of the screen; bottom of the screen), participants did not see one image (e.g., the hawk) over the other part of the image (e.g., the goose). These results are inconsistent with results found in the literature (Dils & Boroditsky, 2010) which suggest that participants should see the image facing in the direction congruent with the spatial location cue (e.g., UPWARD). Also, the results of previous studies (Hommel et al., 2001; Estes et al., 2008) suggest that participant's attention should have been influenced by the spatial location cue shown (e.g., UPWARD).

Hypothesis 1b: Main Effect: Image Location: Ambiguous stimulus: Interpretation

It was expected that there would be a main effect between image location and the interpretation of the ambiguous stimulus. The model used for the logistic regression showed (Table 1) that there was no significant main effect for image location on the interpretation of the ambiguous image. Meaning that, contrary to image location on the screen (e.g., top of the screen) participants did not interpret the image as more hawk-like than goose-like.

This result is inconsistent with the literature, specifically Dils and Boroditsky, (2010) and Stanfield and Zwaan (2001), which showed that participants should see the image that matched the orientation of the spatial location cue (e.g., UPWARD) shown before. The results of the current study do not show these results and did not see the image consistent with the spatial location cue. While the data is not significant, the pattern of results seems to be consistent with the pattern of results found in Chen and Scholl (2014).

Hypothesis 1c: Main Effect: Spatial Language Cue: Ambiguous stimulus: Interpretation

It was expected that there would be a main effect between the spatial language cue and the interpretation of the ambiguous stimulus. The results of the model used for the logistic regression (Table 1) indicated that there was no main effect for spatial language cue and the interpretation of the ambiguous image.

Results indicate that the spatial location cue (e.g., UPWARD) had no bearing on the interpretation of the ambiguous image. Participants were just as likely to interpret the image as both a hawk and a goose after seeing the spatial location cue (e.g., UPWARD). This is inconsistent with the results shown in Dils and Boroditsky (2010) which showed that participants who were given the word “up,” saw the image facing upward when shown an ambiguous image.

Hypothesis 2a: Interaction: Image location X Spatial Language Cue: RT

It was hypothesized that there would be an interaction between image location and spatial language cues on reaction time. Specifically, conditions of congruency were expected to result in quicker reaction times compared to incongruent conditions However, the results of the regression model indicated that there was no significant interaction between image location and spatial location cue when conditions were congruent versus incongruent (d = 0.35, p = .73); R^2 = .004, F(3, 25) = 0.12, p = .73.

Discussion

Previous literature (Estes et al., 2008; Estes et al., 2015; Zwaan & Yaxley, 2003) suggests that certain words (e.g., object words, language cues) can have an effect on attention, influencing the way we perceive an ambiguous stimulus. This research attempted to find the same influence using spatial language cues and an ambiguous image (Dils & Boroditsky, 2010). We expected to find several main effects and interactions between image location, spatial location cue and interpretation, which would have been consistent with previous research (Dils & Boroditsky, 2010 Estes et al., 2008; Estes et al., 2015; Zwaan and Yaxley, 2003).

Contrary to our hypotheses in the study, none of the results were significant at .05 level. With respect to the significance of my finding, only the main effect of image location came close to the threshold for significance. However, the direction of the data shown in the main effect was opposite of what was hypothesized. These patterns of results show more support for the results indicated in Chen and Scholl (2014) than the results of Dils and Boroditsky (2010). Chen and Scholl (2014) suggest that frames and boundaries have an effect on interpretation of an ambiguous stimuli. Such that, participants will favor the image facing inward over the image facing outward. It is in this we see this pattern emerging when looking at the direction of the data with respect to the main effect.

Limitations and Future Directions

There are several limitations, both internal and external to the experiment itself that should be addressed. First and foremost, only one stimulus was used. The data only show the result of one item. This brings two issues: one, we cannot see any variability (if any) within a person's responses. With this, we cannot see if the result(s) would change depending on different items, or the complexity of the image and two: generalizability of findings. Because this research can only generalize to the specific combination of words and image used, future research should also look at expanding both the stimulus used and the number of object words used. Thus, we are not clear that the results are due to the proposed components of the study (e.g., spatial language cues & ambiguous image) or whether it is attributed to the stimulus itself, or a combination.

Future research should look at expanding the pool of both the words and images to test other combinations to see if they have the same (or null) effect. Because of the limited pool of images and words (Dils and Boroditsky, 2010), it would be beneficial to expand on the limited pool of images and words to see if the addition of words and images could elicit the same previously demonstrated effects. Future research could also look at increasing the complexity of the images used as well, to test if any effect occurs when complexity of the images increases. Previous research with the necker cube, an image that changes orientation depending on how you view it, has shown that depending on where you view the cube (as measured via eye-tracking) the perception changes (Einhäuser et al., 2004). Because of these results, one could argue that depending on where you first view an image, the perception might be different as well. If the image is overly complex in nature, the image being perceived could change drastically depending on where participants chose to look first.

The words in the current study, “UPWARD” and “DOWNWARD” were simple in nature and understanding. By increasing the complexity of the given words, it could be seen that the effect demonstrated could dissipate or change. In previous research (Dils and Boroditsky, 2010) abstract language was used in order to see if abstract language could influence perception. However, it was found that under those circumstances that abstract language had no effect (Dils and Boroditsky). Under the paradigm of the current study, it’s worth investigating whether abstract language could be used to potentially see an effect under this paradigm. Also, other research (Estes et al., 2008; Estes et al., 2015; Zwaan and Yaxley, 2003) has shown that words involving items (e.g., shoe, roof, cloud) has shown the same effect as spatial language cues (e.g., UPWARD and DOWNWARD). Arguably, these object words (e.g., roof, shoe) are more complex than spatial language cues. Under a different paradigm, it should be researched to see if these words can elicit the same effect as the current study proposed.

On the note of the stimulus used, only the words “UPWARD” and “DOWNWARD” were used. Because of this, we are unaware if the words “LEFT” and “RIGHT” would have the same effect or work in this same scenario. According to Hommel et al. (2011), results of their study showed that when object location was congruent with the overlearned communication cue, participants were faster to react compared to when the object location was incongruent with the overlearned communication cue (Hommel et al., 2011). Based on these findings, if the overlearned communication cue (or language cue) was placed on the right or left side of the screen, it could be that the effect could still be the same (e.g., participants would look upward) but the time taken to look upward or see the image facing upward could be increased, depending on how long participants had to locate the cue before the image. Also, similar results are shown in Estes et al. (2008) study, with only an interference effect being the difference between the results of Hommel et al. (2011) and Estes et al. (2008). While Estes et al. (2008) did not place objects in the right or left area of viewing, or provide object words (e.g., shoe, hat) that would potentially pull attention to the left or right quadrant of view, it is possible that based off their results and Hommel et al. (2011) that if object words were placed on right of left of view, that the effects would still be the same, albeit slower than if they were placed in the middle. While Hommel et al. (2011) did not explicitly put their overlearned communication cues to the left or right of a subjects view, they did use words such as “LEFT” and “RIGHT” in the experiment which did draw participants' gaze in these directions. It was found that, under these conditions (e.g., “LEFT” and “RIGHT”) that when participants located a target object that matched the direction of the overlearned communication cue, response time was faster in comparison to the condition where the overlearned communication cue (e.g., “LEFT” and “RIGHT”) did not match the location of the target object.

There is also speculation that when using the words “LEFT” and “RIGHT,” the perception of the object might be less consistent because of orientation of the object (e.g., image facing upward and downward) with respect to reaction time. According to Stanfield and Zwaan (2001) orientation of an object mattered when participants were primed with orientation beforehand. Their results show that participants were faster to react to objects that were not congruent with the orientation primed in the sentence (e.g., upright hammer) that was read before the object was shown (e.g., an upright hammer) (Stanfield and Zwaan, 2001). This suggest that if participants were shown the words “LEFT” or “RIGHT” and then an image facing upward, future research could see a drop in reaction time because of the incongruence between the spatial location cue (e.g., “LEFT” or “RIGHT”) and the orientation of the image (e.g., an upward facing image appearing to the left or right).

According to Chen and Scholl (2014) there appears to be an inward bias effect for seeing images facing inward (e.g., toward the frame) over seeing the images facing outward (e.g., away from the frame). While the results were not significant, the pattern seen in the data appears to be the same pattern of results found in Chen and Scholl (2014). The frame of the computer could be influencing the attention (e.g., the direction view the image) of the participants more so than the spatial location cue seen before. Also, due to limitations in the programming of the experiment, there was no true counterbalancing of the participants. Participants were placed in conditions completely at random. Thus, more participants could have been placed in the condition showing

"UPWARD" and having the image in the upward location of the screen, compared to the condition of having "DOWNWARD" and having the image placed at the bottom of the screen. Lastly, the overall sample for the experiment was lower than anticipated (e.g., 330 participants needed to reach sufficient power). Because of this, the results do not reflect what could have been if the experiment had been able to collect the total number of participants needed.

Another major limitation of this thesis was the pandemic that occurred, COVID-19. Because of the pandemic, the experiment was created online, which created some additional limitations. The experimenter was not able to moderate and account for external factors or conditions (e.g., participants being on their phone, getting distracted, noises). Also, the device and screen size that the experiment was taken on was not consistent and could not be controlled. This could have an effect on the results due to the variations of personal computer screen monitors (e.g., size, shape, orientation, etc.). While the data was able to be removed, "bots" were also an issue in the initial data collection process. Bots are an issue for all online experiments and can create a problem in the data analysis. However, the data that were collected was able to sub select "bots" out of the data and have their data removed.

Future research in this area should take note of the pandemic that the current study took place in. The pandemic created a unique situation for the current study that future research should take note of when attempting to extend or replicate the current study. If the experiment is to be replicated, the replication study should aim to conduct the experiment in a physical environment where participants can be monitored and moderated to control for external factors (e.g., distracted participants, being on their phones, noise, screen size). Also, future research should attempt to create the experiment on a more flexible program, to better control of items such as the ability to counter-balance.

Implications

With respect to perception and ambiguous imagery, very little is known about what specifically affects perception of ambiguous imagery in certain ways. The results of this study add onto the literature of language effects on perception of ambiguous imagery, showing the effects of language cues (e.g., UPWARD and DOWNWARD) and their effects on attention and by extension perception of ambiguous images. Implications of this study also include being more critical about how items are presented to people, if you wish to avoid a certain outcome or bias. If you have a situation where you do not want someone to be influenced by any factors or stimulus, this research shows that perhaps you should avoid showing participants any type of language cue(s) before showing them the image you want them to see. However, the opposite could also be an implication as well, in that if you want to force or influence someone to see something in a specific manner, it might be worth showing them some type of language stimulus (e.g., “UPWARD”) before showing them the image you want them to interpret. In doing so, you influence what the person might see first, getting results or outcomes in your favor that you want.

While the study did not find significant results, it is worth noting that participants may have been influenced by their environment (e.g., the frame), just not the influence of our spatial location cue. In this, we see that perception of an ambiguous image is influenced by objects (i.e., the frame of the computer) that are not being manipulated but are standing alone. It is possible that the influence of the frame superseded the influence of the spatial location cue, as participants would have seen the frame after the cue had disappeared. In Dils and Boroditsky (2010), the image used was not located near the frame of the computer, thus participants may have avoided any influence. This could explain the difference in results between the current study and Dils & Boroditsky (2010).

Conclusions

The current study tested the relationship between spatial language cues, attention and ambiguous stimulus interpretation. Through the use of spatial language cues (e.g., “UPWARD” and “DOWNWARD”) and an unrelated ambiguous image (e.g., goose/hawk image), the hope was to show a causal relationship between the language cues and interpretation of the ambiguous stimulus. Dils and Boroditsky (2010) have shown that unrelated spatial language cues can have an effect on interpretation of an ambiguous image (e.g., goose/hawk image). Estes et al. (2008) results also show that object words (e.g., hat, shoe) are enough to draw our attention to specific areas in our environment. Language has also been shown to help us identify objects in our environment faster (Estes, Verges & Barsalou, 2008).

Upon completing the analysis of the results, none of the findings support the literature. The spatial language cues did not have a significant effect on attention to influence the interpretation of the ambiguous image. This is in contrast to what was found in Estes et al. (2008), which showed object words (e.g., shoe) could influence participants' attention. Also, contrary to results found in Dils and Boroditsky (2010), results were not consistent, and participants did not see the image that was congruent with the spatial location cue (e.g., “UPWARD” or “DOWNWARD”). Lastly, results indicated that the location of the image did not matter, nor did the word shown to participants. This is in contrast to the literature that shows that both items should have an effect on interpretation of an ambiguous image (Dils & Boroditsky, 2010; Estes et al, 2008; Estes et al., 2015).

While the results do not support the literature, the study contributes to the ever-growing literature on spatial language cues (e.g., UPWARD), ambiguous images (e.g., the goose/hawk image) and attention. What we see in our environment could be perceived differently based on what we have previously seen or what is actively drawing our attention in the current moment. If someone is attempting to make a judgement call or come to a conclusion based on an interpretation, they should take a second to think about the environment and the influences it might have on their perception in a situation that is open to interpretation or ambiguous in nature.

References

Andrews, S., & Bond, R. (2009). Lexical expertise and reading skill: Bottom-up and top-down processing of lexical ambiguity. *Reading and Writing: An Interdisciplinary Journal*, *22*(6), 687–711. https://doi.org/10.1007/s11145-008-9137-7

Bernstein, L., & Cooper, L. (1997). Direction of motion influences perceptual identification of ambiguous figures. *Journal of Experimental Psychology: Human Perception and Performance*, *23(*3*)*, 721–737. https://doi.org/10.1037/0096-1523.23.3.721

Chen, Y.-C., & Scholl, B. (2014). Seeing and liking: Biased perception of ambiguous figures consistent with the “inward bias” in aesthetic preferences. *Psychonomic Bulletin & Review*, *21*(6), 1444–1451. https://doi.org/10.3758/s13423-014-0617-2

Cognition. Run experiments online. Cognition. (n.d.). https://www.cognition.run

Daelli, V., van Rijsbergen, N., & Treves, A. (2010). How recent experience affects the perception of ambiguous objects. *Brain Research*, *1322*, 81–91. https://doi.org/10.1016/j.brainres.2010.01.060

Dils, A. T., & Boroditsky, L. (2010). Processing unrelated language can change what you see. *Psychonomic Bulletin & Review, 17*(6), 882-888. https://doi.org/10.3758/PBR.17.6.882

Dudschig, C., Lachmair, M., de la Vega, I., & Filippis, M. (2012) From top to bottom: Spatial shifts of attention caused by linguistic stimuli. *Cognitive Processing.* 13(Suppl 1): 151. https://doi.org/10.1007/s10339-012-0480-x

Dumitru, M., Joergensen, G., Cruickshank, A., & Altmann, G. T. M. (2013). Language-guided visual processing affects reasoning: The role of referential and spatial anchoring. *Consciousness and Cognition: An International Journal*, *22*(2), 562–571. https://doi.org/10.1016/j.concog.2013.03.008

Einhäuser, W., Martin, K. A. C., & König, P. (2004). Are switches in perception of the necker cube related to eye position? *European Journal of Neuroscience*, *20*(10), 2811–2818. https://doi.org/10.1111/j.1460-9568.2004.03722.x

Estes, Z., Verges, M., & Adelman, J., (2015) Words, objects, and locations: Perceptual matching explains spatial interference and facilitation, *Journal of Memory and Language*, *84*, 167-189, https://doi.org/10.1016/j.jml.2015.06.002

Estes, Z., Verges, M., & Barsalou, L. W. (2008). Head up, foot down: Object words orient attention to the objects' typical location. *Psychological Science*, *19*(2), 93–97. https://doi.org/10.1111/j.1467-9280.2008.02051.x

Glenberg, A., & Kaschak, M. (2002). Grounding language in action. *Psychonomic Bulletin & Review, 9*(3), 558–565. https://doi.org/10.3758/BF03196313

Hommel, B., Pratt, J., Colzato, L., & Godijn, R. (2001). Symbolic control of visual attention. *Psychological Science*, *12*(5), 360–365. https://doi.org/10.1111/1467-9280.00367

Kaschak, M., Zwaan, R., Aveyard, M., &Yaxley, R. (2006), Perception of auditory motion affects language processing. *Cognitive Science*, *30*(4)733-744. https://doi.org/10.1207/s15516709cog0000_54

Kingstone, A., Smilek, D., Ristic, J., Friesen, C., & Eastwood, J. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science*, *12*(5), 176–180. https://doi.org/10.1111/1467-8721.01255

Klemfuss, N., Prinzmetal, W., & Ivry, R. (2012). How does language change perception: A cautionary note. *Frontiers in Psychology*, 3, 78. https://doi.org/10.3389/fpsyg.2012.00078

Kosegarten, J., & Kose, G. (2014). Seeing reversals in ambiguous images: To know or not to know? *Perceptual and Motor Skills*, *119*(1), 228–236. https://doi.org/10.2466/24.27.PMS.119c15z9

Mast, F., & Kosslyn, S. (2002). Visual mental images can be ambiguous: Insights from individual differences in spatial transformation abilities. *Cognition*, *86*(1), 57–70. https://doi.org/10.1016/S0010-0277(02)00137-3

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). On the experiential link between spatial and temporal language. *Cognitive Science*, *29*(4), 655–664. https://doi.org/10.1207/s15516709cog0000_17

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science*, *18*(11), 1007–1013. https://doi.org/10.1111/j.1467-9280.2007.02016.x

Ostarek, M., & Vigliocco, G. (2017). Reading sky and seeing a cloud: On the relevance of events for perceptual simulation. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *43*(4), 579–590. http://dx.doi.org/10.1037/xlm0000318

Palef, S., & Olson, D. (1975). Spatial and verbal rivalry in a Stroop-like task. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, *29*(3), 201-209. https://doi-org.ezproxy.humboldt.edu/10.1037/h0082026

Richardson, D., Spivey, M., Barsalou, L. & McRae, K. (2003), Spatial representations activated during real- time comprehension of verbs. *Cognitive Science, 27*(5), 767-780. https://doi.org/10.1016/S0364-0213(03)00064-8

RStudio Team (2020). *RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.*

R Core Team (2021). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. https://www.R-project.org/.

Signorell, A. (2021). DescTools: Tools for descriptive statistics. R package version 0.99.41. https://cran.r-project.org/package=DescTools.

Stanfield, R., & Zwaan, R. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156. https://doi.org/10.1111/1467-9280.00326

Sobel, K., Pickard, M., & Acklin, W. (2009). Using feature preview to investigate the roles of top–down and bottom–up processing in conjunction search. *Acta Psychologica*, *132*(1), 22–30. https://doi.org/10.1016/j.actpsy.2009.06.003

Wilson, A., & Golonka, S. (2013). Embodied Cognition is not what you think it is. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00058

Zwaan, R., Stanfield, R., & Yaxley, R. (2002). Language comprehenders mentally represent the shape of objects. *Psychological Science*, *13*(2), 168–171. https://doi.org/10.1111/1467-9280.00430

Zwaan, R. & Yaxley, R (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*(4), 954. https://doi.org/10.3758/BF03196557

Appendix