Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for a degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter now, including display on the World Wide Web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Justin Whitney March 25th 2025

Assessing the Effects of Intervening Elements on Nonadjacent Dependency Learning

By

Justin Whitney

Benjamin Wilson, PHD

Advisor

Neuroscience & Behavioral Biology

Benjamin Wilson, PHD

Advisor

Robert Hampton

Committee Member

Keith Easterling

Committee Member

2025

Assessing the Effects of Intervening Elements on Nonadjacent Dependency Learning

By

Justin Whitney

Benjamin Wilson, PHD

Advisor

An abstract of

a thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2025

Abstract

Assessing the Effects of Intervening Elements on Nonadjacent Dependency Learning

By: Justin Whitney

Background: Statistical learning is a key component to humans developing language. Non-human primates can learn through statistical learning too. Statistical regularities that are seen in language are adjacent and non-adjacent dependencies. Non-human primates have been able to demonstrate learning of these statistical regularities. Here we use different sequences to compare if monkeys can learn non-adjacent dependencies of varying lengths.

Methods: We used a serial reaction time paradigm to look at Rhesus Macaques reaction times to stimuli on a computer screen. We created three different sequences, each containing different dependent elements. We tested the monkeys with these sequences in four phases: Baseline, Testing, Generalization, and Baseline 2. The monkeys completed thousands of trials the experiment to allow us to analyze their reaction times.

Results: Initially we concluded that the first monkey showed a decrease in reaction time for the dependent elements, but after more statistical tests we found that the monkey did not learn the dependencies. The second monkey did not show a decrease in reaction time for the adjacent dependency sequence. The last monkey showed learning in the adjacent dependency sequence, but no learning in the non-adjacent dependency sequences.

Conclusion: Learning adjacent and non-adjacent dependencies are difficult in monkeys. One of the monkeys was able to learn the adjacent dependency, but none of the monkeys learned the non-adjacent dependencies.

Assessing the Effects of Intervening Elements on Nonadjacent Dependency Learning

By

Justin Whitney

Ben Wilson, PHD

Advisor

A thesis submitted to the Faculty of Emory College of Arts and Sciences

of Emory University in partial fulfillment

of the requirements of the degree of

Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2025

Table of Contents

Introduction

Language is a key component in how humans communicate with others, and understanding the underlying cognitive processes supporting language is crucial to understanding the different components that allow language to function. In humans, Broca's area, which is involved in speech production, and Wernicke's area, which is involved in comprehension, are critical in allowing humans to have language. These two brain regions are connected by nerve fibers known as the arcuate fasciculus. Non-human primates have homologs to Broca's area known as area 44 and 45 (Gallagher & Zilles 2019), however, language is unique to only humans. Studying other animals with recent common ancestors to humans can allow us to investigate different components of language and provide insights about how language may have evolved. Humans begin to learn language as infants through a process called 'statistical learning'. Statistical learning is the process through which we recognize patterns in environmental stimuli (Saffran et al 1996). For example, by noticing statistical regularities in streams of continuous speech, infants begin to detect word boundaries (identifying where words start and end; Saffran et al., 1996), and even learn simple grammatical rules (Gomez et al., 1999). In addition to humans, other animals, including non-human primates, can also learn patterns through statistical learning (Santolin et al 2018). Therefore, understanding the limits of statistical learning in non-human primates can give us a better understanding of the evolution of cognitive processes that are required for human language.

There are many different types of learning that are unique from statistical learning, such as classical conditioning and operant conditioning. These types of conditioning are reinforced. For example, in classical conditioning a dog can associate the ring of a bell with a certain food

reward like in Pavlov's dog experiment. In operant conditioning a rat can associate hitting a lever with a food reward. Both of these types of learning are reinforcement, whereas in statistical learning reinforcement with a reward or punishment does not occur.

Much of the research on statistical learning has focused on the types of dependencies that animals (or infants) can learn (Wilson et al 2020). A very simple type of statistical relationship is known as an adjacent relationship or adjacent dependency, in which one stimulus (e.g., 'A') predicts the stimulus that immediately follows it in a sequence (e.g., 'B', in the sequence 'AB'). In English, adjacent dependencies are common, for example in tense agreement and subject verb agreement (Wilson et al. 2020). An example of an adjacent dependency in subject verb agreement is "the cat jumps" or "the cats jump." In both cases, one needs to ensure that there is plural agreements between the noun and the verb, which appear adjacent to one another in the sentence. However, beyond adjacent dependencies, language also contains many longer distance, 'non-adjacent' dependencies. These are a more complex type of dependency (Wilson et al., 2020), in which the same 'A' and 'B' elements are separated by one or more intervening elements (eg., 'AxB', 'AxxB', etc.). In non-adjacent dependencies, the B element is still dependent on the A element, but the A element must be held in working memory until the B element is encountered. Returning to our linguistic example, non-adjacent dependencies in plural agreement could be seen in the phrases: "the cat [on the grass] jumps" or "the cats [on the grass] jump." Again, one must track the plural agreement between noun and verb, but now these are separated by an intervening phrase, creating a non-adjacent dependency. Processing these non-adjacent dependencies requires keeping track of the initial 'A' element for a longer time, adding additional cognitive and working memory demands compared to processing adjacent dependencies (Wilson et al., 2020).

Both humans and non-human primates can learn adjacent dependencies (Fitch & Hauser, 2004; Gebhart, Newport, & Aslin, 2009; Pacton, Sobaco, & Perruchet, 2015; Reber, 1967; Saffran et al., 2017; Wilson, Smith, & Petkov, 2015). However, while studies have reported that non-human primates may learn nonadjacent dependencies (tamarins, Versace et al., 2017; baboons, Malassis, Rey, & Fagot, 2018), these results are much more variable suggesting that learning nonadjacent dependencies might be more difficult than learning adjacent dependencies for nonhuman primates (Wilson et al., 2020). Understanding the limitations on monkeys' abilities to learn non-adjacent dependencies is crucial for understanding the cognitive processes that allowed language to evolve in humans.

Different experimental paradigms have been used to investigate statistical learning in non-human primates such as Artificial Grammar Learning paradigms (Saffran, Aslin, & Newport, 1996) and Serial Reaction Time paradigms (Nissen & Bullemer, 1987). In these experiments, the monkeys are presented with sequences of stimuli that have been artificially generated by a set of rules (an artificial grammar) that dictates the ordering of elements in a sequence.

In artificial grammar learning experiments, monkeys are first exposed to these sequences and then are tested on their ability to distinguish between sequences of stimuli that follow and do not follow the artificial grammar rules. In these tasks, stimuli are typically auditory sequences presented from speakers, and monkeys' responses are assessed in terms of orienting responses towards the audio speaker (Wilson et al., 2013). If monkeys show stronger dishabituation orienting responses when ungrammatical sequences are presented relative to grammatical sequences, this demonstrates learning.

In Serial Reaction time tasks, rather than presenting sequences of auditory or visual stimuli, participants are presented with a sequence of stimuli in differential spatial locations (e.g., on a

touchscreen monitor), and they are required to touch each stimulus as soon as it appears. The spatial locations in which the stimuli occur is determined by the rules of an artificial grammar, and this produces statistical regularities such that some stimulus locations predict the location of upcoming stimuli. If participants are sensitive to these statistical regularities, they should be able to respond to these predictable stimulus locations more quickly, and this decrease in reaction time is taken as evidence of learning.

The goal of the current experiment is to use a serial reaction time paradigm to determine if rhesus macaques can learn non-adjacent dependencies of varying lengths. This will be done by creating three different artificial grammars that generate three different dependencies, each containing elements that are predictable. We predicted that if the monkey's learn these dependencies, then they will show a decrease in reaction time.

Methods

The goal of this project was to assess whether monkeys learn dependencies of different distances using a serial reaction time paradigm. There are two experiments in this project and the general methods for both of them are similar. I will first describe the general methods and then provide more specifics for each experiment in turn.

General methods

In a serial reaction time task, a series of visual stimuli were presented sequentially on a touchscreen computer monitor, in different spatial locations. We presented the participants an array of outlined white circles. One of the circles was filled in red and the participants had previously been trained to touch these colored stimuli (See fig 1.). The stimuli could either appear in random locations (as in the Baseline phase, see Procedure, below), or the order of the

locations in which the stimuli appeared could be governed by an artificial grammar (as in the Testing phases, and see Stimuli, below). The artificial grammar produces statistical regularities in the sequences of spatial locations, potentially allowing participants to predict where upcoming stimuli may appear. Learning is evident if the participants respond faster to these predictable sequences compared to the unpredictable sequences.

We tested rhesus macaques using this serial reaction task paradigm. Macaques were individually tested in their home cages, using custom designed cognitive testing systems, consisting of a touchscreen computer and a food reward dispenser. The testing systems were available to the monkeys for approximately 6 hours per day, and the animals were free to engage with them at any time.

Procedure

In both experiments, trials began with the presentation of a 'go' stimulus, which the monkeys touched to start the trial. They were then presented with a 4x3 array of empty white circles on a black background (Fig, 1). We then presented sequences of stimuli (see Stimuli, below), in the form of red circles which appeared at different spatial locations. The monkeys were trained to touch these red circles as soon as they appeared. The monkeys received a food pellet as a reward if they responded to all of the stimuli in the sequence correctly. However, the trial was aborted as soon as the monkey made an incorrect response by touching the wrong stimulus. The aborted sequences were presented again to the monkeys after they were presented with all the sequences.

Experiment 1, Methods

The goal of the first experiment was to simultaneously assess the learning of adjacent and nonadjacent dependencies in rhesus macaques.

One female rhesus macaque (age 5), Cassie, took part in this experiment.

Procedure

The monkeys took part in four different phases of the experiment: Baseline (with random sequences), Testing (containing a subset of sequences generated by the artificial grammar), Generalization (containing novel grammatical sequences the monkey had no seen before), and a final Baseline phase (containing random sequences).

Phase 1: Baseline. First, we conducted a baseline phase to familiarize the monkey with the SRT procedure, and to allow us to acquire baseline reaction time data which showed how quickly the monkey tended to respond to random, unstructured sequences. During the baseline phase of the experiment, we presented random sequences of stimuli that did not conform to any artificial grammar. The sequences were generated by randomly selecting stimulus positions with the only constraint being that the same position could not be repeated in a sequence. Each sequence contained 6 elements (randomly selected spatial locations), and we presented 1000 trials during this baseline phase.

In the Baseline phase, none of the stimulus locations are predictable based on prior stimuli in the sequences, thus we predicted no differences in reaction times between any of the stimulus locations.

Phase 2: Testing. In the second phase of the experiment the monkey was exposed to sequences containing predictable dependencies. For the purposes of this task, we created predictable

relationships between some of the spatial locations that stimuli could occupy, such that three of the locations (labelled A_1, A_2 and A_3, Fig 2) were always followed by stimuli in a specific location (B_1, B_2 and B_3, respectively, Fig 2.). As mentioned earlier, the B elements do not necessarily directly follow the A elements but will appear later in the sequence, either as an adjacent or non-adjacent dependency. The other six locations were labelled as x_1-x_6, and were not predictable or predictive based on the A and B stimulus locations(Fig. 2).

In this experiment, we used three types of sequences, which were intermixed during testing. The first of these contained an adjacent dependency and took the form xxA_1B_1xx (where any of the x elements could be used in place of each of the xs in the sequence, See Fig. 3A). The second type of sequence contained a nonadjacent dependency with one intervening x element (i.e., xA_2xB_2xx, See Fig. 3B), and the final type of sequence contained a longer nonadjacent dependency with two intervening x elements (i.e., xA_3xxB_3x, See Fig 3C). Note that different pairs of A and B elements were used in the three different types of sequence, so the presence of any given A (e.g., A_1) could predict both which B element would follow (i.e., B_1) and also how long the dependency between the A and B elements would be (in this case, an adjacent dependency with B_1 immediately following A_1). These sequences occurred evenly in this phase, but occurred in a random order, with the sequence types being intermixed.

There are a total of 360 possible variations of each sequence type (a total of 1080 for all 3 sequence types). During the testing phase, we used 240 sequences of each sequence type, reserving 120 of each type for the subsequent generalization phase. In this Testing phase, the monkey completed 2880 trials, with 960 trials per sequence type.

We predict that if learning occurs, we would see a decrease in the reaction times for the B elements in comparison to the B elements' reaction time in the baseline phase. Additionally, we

predict that as the number of x elements in between the A and B elements increases that the B element reaction time will decrease less from baseline. The longer the dependencies, the longer the learning will take and the less the reaction time will decrease. We also predict that as the monkey completes more trials, she will have a faster reaction time for all the elements because she is learning the SRT task, but she will have a larger drop for the B elements because they are predictable from the A element.

Phase 3: Generalization. The third phase of this experiment was the generalization phase, which tested if the monkey would generalize the previously learned information from the Testing phase to novel sequences they had not experienced before. In the prior Testing phase, we withheld 120 variations of each sequence type so that we could present those variations in the generalization phase. We presented 1080 trials, 360 per sequence type, for the monkey in this phase.

If the monkeys had learned the dependencies between the A and B stimulus locations, then we would predict that this would generalize to the novel sequences used in the Generalization phase, and would continue to show faster reaction times to the B elements. However, if the monkeys were instead responding based on their familiarity with the sequences used during the testing (e.g., by memorizing particular sequences), then we would expect slower reaction times to the B elements compared to the testing phase. Note that the Testing phase contained 720 unique sequences, so memorization seems unlikely, but assessing generalization allows us to confirm dependency learning separate from familiarity with specific stimulus sequences.

Phase 4: Baseline 2. The final phase of the experiment was a second baseline phase using random sequences, and was identical to the baseline phase described above. The purpose of having another baseline phase after the testing and generalization phase was to see if there was a

reaction time decrease due to learning or due to getting faster at the reaction time task. It is possible that decreases in reaction time might be attributable to practice on the task, rather than dependency learning. Therefore, we predict that if the monkey learned the dependencies, we would see an increase in reaction times to the 4th element of the sequences in this baseline phase (which is the position in the sequence in which the B stimuli previously occurred) relative to the Testing and Generalization phases. Conversely, if we see no differences in responses to the 4th sequence element between Testing and/or Generalization and this Baseline phase, this could suggest a general increase in speed, rather than dependency learning.

Data Analysis:

First, we removed all of the outlying reaction times. An outlier was defined as being more than 3 standard deviations from the mean reaction time per transition or a reaction time greater than 5 seconds, or shorter than 200ms. 11.3% of the total trials were removed as outliers.

The data analysis was done by calculating the reaction time to each element and analyzing whether there was a larger decrease in the reaction time to the B element(4th element), compared to the other elements.

Additionally, we compared these reaction times in all four phases of the experiment(Baseline 1, Testing, Generalization, Baseline 2). As we only had a single participant, standard statistical tests (e.g., t-tests) are inappropriate as they assume independence between datapoints. Therefore, we used permutation tests for statistical comparisons, as these have fewer assumptions and allow within subject comparisons. All analyses were conducted in Matlab.

Experiment 1, Results

In this experiment our goal was to test non-adjacent dependency learning using a serial reaction time paradigm in monkeys. We predicted that, if the monkey's had learned a dependency (for example the adjacent, A_1-B_1 dependency), then they should respond faster to the predictable B element compared to the less predictable A and x elements. Similarly, if monkeys learned the longer distance relationships between A_2 and B_2 or A_3 and B_3, we would again predict faster reaction times to the predictable B elements. Critically, in this experiment although the distance of the A-B dependencies varied, the B element was always the 4th element of the 6 element long sequences (i.e., xxA_1B_1xx, xA_2XB_2xx and A_3xB_3xx). Therefore, we compared the reaction times to the 4th element of the sequences to the other elements of the sequence. We conducted these analyses both in the Baseline phases (in which the sequences were random, therefore we should expect no difference in reaction time between the 4th element and any other element of the sequence), as well as in the Testing and Generalization phases (where we would predict these differences if learning had occurred).

In the first baseline phase, as predicted, there was no difference in reaction time between the 4th element of the sequence and the other elements, besides the first element. This monkey was slow at responding to the first element across all the phases of this experiment regardless of the spatial location that it was located in. This was a prediction for this experiment, and the first element is not particularly less predictable than the other x elements in the sequences. Rather this appears to represent an idiosyncrasy of this specific macaque, who consistently responded slowly to the first element of the sequence, before increasing to a more consistent speed across other elements.

We ran a series of permutation tests (See Table 1), demonstrating that the monkey responded to all of the elements in the sequence, besides the first element, with similar speeds. In the Testing phase, we found that the monkey responded significantly faster to the B elements than to the other elements in the sequence (see Table 1). This suggests that the monkey had learned the dependent B elements were predictable. In the subsequent generalization phase, we found the same results with faster reaction times to the B elements, which demonstrates that this increase in reaction speed is due to learning the A-B dependencies rather than memorizing all of the specific sequences used during the testing phase(Table 1). Finally, in the last Baseline phase, we found that the 4^{th} element only converged with one of the other elements which demonstrates that she may have just gotten faster at the 4^{th} element in the sequence, rather than the dependent B element (Table 1).

The prior analyses showed that the monkey responded faster to the B elements than other elements in the sequences in the Testing and Generalization phase (but not the 4^{th} element of the sequence in the first Baseline phase), as predicted if the monkey had learned the dependency between the A and B elements.

We next analyzed the reaction times separately for the three different types of sequences, those containing adjacent dependencies (xxA_1B_1xx) and those containing nonadjacent dependencies (A_2xB_2xx and A_3xxB_3xx). We had predicted that adjacent dependencies would be easier to learn than nonadjacent dependencies, and thus that we would see the largest decrease in reaction times, in the testing phase, to the B element in the adjacent sequences, compared to the nonadjacent sequences. However, we found no differences between the different sequence types (Fig. 4, Table 2).

This suggests that the monkey simultaneously learned all three types of dependencies, and performed similarly across these different sequence types, despite our prediction that the nonadjacent dependencies should be more difficult to learn (see Wilson et al., 2020).

We predicted that the reaction times would increase for the 4th element in the sequences in the Baseline 2 phase, when the sequences became unpredictable again. However, as noted above, we found that responses to the 4th element remained faster than to most other elements (Table 1; Fig., 5.) This difference, except for the first element, was on average 23ms, which is very minimal compared to other significant differences we observed, but this raises the question of if Cassie had learned the dependent B element. We also noted that the effect that we saw in the testing phase (with B elements eliciting faster responses than the other elements) occurred immediately, from the first block of trials in the phase(See Fig 5). This is surprising, as learning was predicted to occur gradually, over the course of the experiment. Cassie's reaction times for the first 400 trials and last 400 trials in the testing phase were statistically the same (P value of .3476). Therefore, we conducted follow up analyses, to assess whether the faster reaction times to the B elements in the 4th position of the sequences might instead be due not to learning the dependencies, but simply that these stimuli appeared in spatial locations in which the monkey happened to respond very quickly.

To test this, we compared reaction times to the locations of the B elements during the testing phase with the same spatial locations (which is distinct from the same ordinal position in the stimulus sequences, which we did in the prior analyses) in the Baseline phases. Here, we found that the reaction times for this B position were statistically the same across all phases (Table 3; Fig., 4). This indicates that rather than learning the A-B dependencies, the differences

in reaction times that we saw might instead be attributable to the spatial locations selected for the B elements in the Testing and Generalization phases.

Discussion

The results of Experiment 1 originally appeared to show that learning had occurred. The monkey showed a decrease in the B reaction time in the Testing and Generalization phase compared to the Baseline phase. At first this would appear to indicate learning, because as predicted, we observed a significant decrease in reaction time in the Testing and Generalization phases compared to the Baseline phase. However, as noted above, there are alternative explanations, and learning does not appear to have occurred. The monkey immediately showed a decrease in reaction times in the first block of 200 trials of the testing phase. In these experiments, learning typically occurs overtime, not immediately. Additionally, Cassie's data showed that she responded equally quickly to all of the B elements (B1, B2, and B3, in both adjacent and non-adjacent conditions) at the same rate, which would be surprising because previous literature has shown that adjacent dependencies are easier to learn than non-adjacent dependencies.

Given these unexpected patterns in the data, we considered other possible explanations for the faster reactions we observed to the B elements. The initial analyses compared the reaction times to the 4^{th} element in a sequence (which was always a B in the Testing and Generalization phase), to the 4^{th} element in the sequence in the baseline phases. While this seemed like a reasonable approach, this meant that we were comparing reaction times to different spatial locations between the Baseline and Testing/Generalization phases. When we compared the reaction times to the same spatial locations across phases, we found no differences. Therefore, it

appears that, rather than learning the dependencies, an experimental confound led to the apparent decreases in reaction time to the B elements in the Testing and Generalization phase.

The stimulus positions for the Testing and the Generalization phases were selected based on Cassie's reaction times in the initial Baseline phase. The positions that had the least variance between all possible transitions between elements were chosen as the final positions for the Testing and Generalization phases. This is how the B element position was selected. To avoid a problem like this in the future, we could calculate the B element position in a different way, for example selecting the location with the average reaction time compared to all 12 locations.

Experiment 1 led us to design a second experiment where we separated the sequence types so that the monkeys only saw one sequence type at a time. We hoped that this would help aid their learning because instead of having to learn three patterns simultaneously, they will only have to focus on one sequence type.

Experiment 2

The goal of the second experiment was to independently assess the learning of adjacent and nonadjacent dependencies in rhesus macaques.

Participants

Two female rhesus macaques (age 5), Cleo and Zoe, participated in this experiment.

Stimuli

We followed a similar experimental setup as in experiment 1. This experiment contained similar phases as experiment one, but instead of combining all of the sequences into one phase, we

separated them (see below). The aim of this was so that instead of the monkeys focusing on 3 different sequences simultaneously, they would only have to focus on one at a time. Additionally, instead of 6 elements in a sequence, we reduced the elements to 5. This was because the 6th element was not important in our data collection and the monkeys could do more trials with fewer elements per sequence. This reduced the number of variations per sequence to 120.

In experiment 1 all of the elements were presented as red circles. However, here we wanted to emphasize the specific dependent elements and teach the monkeys that they are related, so we presented the A and B elements as blue circles. The X elements were presented in red, as in Experiment 1.

Phase 1: Baseline. This phase was the exact same as Experiment 1, but with 5 elements in a trial.

Phase 2: Testing. In this phase, we tested the monkeys on each sequence type individually. The rationale behind introducing the monkeys the sequences individually was that we wanted to make it easier for the monkeys to learn the dependencies and learning one dependency at a time may be easier than learning three simultaneously. Since there was a reduction of elements in the sequence from 6 to 5, there was also a reduction in the number of variations per sequence type. There were 120 variations of each sequence type, and we removed 40 of them from the Testing phase to be presented to the monkeys in the subsequent generalization phase. The monkeys participated in 2400 trials for each sequence type, 30 repetitions of each individual sequence.

Phase 3: Generalization. As in Experiment 1, the Testing phase was followed by a Generalization phase. We presented the 40 sequences that were not presented during the testing phase 5 times, in addition to all the other sequences, for a total of 600 trials.

Unlike Experiment 1, in Experiment 2 there were 3 different Testing and Generalization phases (for each of the three sequence types, xxA_1B_1x, xA_2xB_2x, and A_3xxB_3x, Fig. 2). Once the monkey completed both the testing and generalization phases for a given sequence type, they moved onto the next sequence type.

Phase 4: Baseline 2. This phase was the same as the first Baseline phase.

Data Analysis

The data analysis was the same as experiment 1. We removed the outliers, as defined in experiment 1, and found that in Cleo's data 7.84% of the trials were outliers and in Zoe's data 7.64% of the trials were outliers.

Results experiment 2

The goal of this new experiment was to present the monkeys with sequences containing adjacent and non-adjacent dependencies if separate phases, to assess learning. In this experiment the B elements (which always appear in the 4^{th} position of each sequence) can be predicted based on the preceding A element. However, it is also possible to learn that the 4^{th} element of the sequence always appears in the same location (for any given sequence type), independent of learning the relationship to the A element. Similarly, because we only present a single sequence type in each phase, the A element also always occurs in a fixed spatial location and a fixed position in the sequence (the third position in xxA_1B_1x, the second position in xA_2xB_2x, and the first position in A_3xxB_3x). Therefore, both the A and B stimuli can be predicted based on learning that they occur in these fixed locations and ordinal positions in the sequence. Therefore, if the monkeys learn this information, we would predict that reaction times would decrease for both the A and the B

elements in the Testing and Generalization phase, relative to the Baseline phases. If the monkeys have additionally learned that the A element predicts the B element, we would also predict that the reaction times to the B elements would be even faster than to the A elements.

Regarding the different sequence types, we predicted that, if the monkeys learned the adjacent dependency (A_1-B_1), then they should respond faster to the predictable B_1 element compared to the less predictable A and X elements. Similarly, if monkeys learned the longer distance relationships between A_2 and B_2 or A_3 and B_3, we would again predict faster reaction times to the predictable B elements. Critically, in this experiment although the distance of the A-B dependencies varied, the B element was always the 4th element of the 5 element long sequences (i.e., XXA_1B_1X, XA_2XB_2X and A_3XXB_3X). Therefore, we compared the reaction times to the 4th element of the sequences to the other elements of the sequence. We conducted these analyses both in the Baseline phases (in which the sequences were random, therefore we should expect no difference in reaction time between the 4th element and any other element of the sequence), as well as in the Testing and Generalization phases (where we would predict these differences if learning had occurred).

In Experiment 2, two monkeys were tested. The first monkey, Cleo, only completed the adjacent dependency condition, as she failed to show any learning even of the adjacent dependencies. Given that we predicted that the non-adjacent dependencies would be harder to learn, we did not advance her to these phases of the Experiment. During the adjacent dependency condition, her reaction times to all of the different elements were not statistically different. (See table 4)(Fig 6). This indicates that a change in reaction time in the A or B element in the testing phase could be from learning.

There was also no change in reaction time for the A or B elements between the testing or generalization phase compared to the baseline phases(Table 5, Fig 7).This indicates that the monkey did not learn the dependency between the A or B element, or to predict these elements based on their appearance in the fixed positions in the xxA_1B_1x sequence. Therefore, we decided to take her off the experiment.

The second monkey, Zoe completed all of the three sequence types. In her baseline phase, there was no difference in the reaction times between most of the elements, as predicted(table 6). However, the monkey did respond much slower to the first element of the sequence. This was not predicted, but as the critical element in the sequence is the B element in the 4th position, it is still possible to conduct further analyses, despite the monkey's initial slowness on each trial.

We analyzed each sequence type separately, to see if there were any decreases in the A and B element reaction times. First, we ran tests to see if there were any decreases in reaction times across the Testing phases for each sequence type. Learning does not occur immediately so we predict that the start of the testing phase reaction times for the A and B element will be greater than the end of the testing phase reaction times for the A and B elements. In the adjacent dependency sequence (xxABx) she showed a decline in reaction times for both the A and B elements when comparing the first 500 trials to the last 500 trials in the testing phase ($p = .0055$ and $p < 0.0001$, respectively). This indicates that she is showed learning of the A and B elements, although more analysis is needed to determine the type of learning that is being done(see below). This decrease in reaction time is continued in the generalization phase indicating that she was able to apply her learning to new stimuli ($p = 0.0051$ and $p = 0.0026$, respectively).

In the xAxBx sequence testing phase, the reaction times for both the A and B elements did not change from the first 500 trials of the testing phase trial to the last 500 trials of the testing phase ($p = .3800$ and $p = .1358$, respectively). However, there was a decrease in reaction time for the B element from the first 500 trials in the testing phase to the last 500 trials in the generalization phase ($P = .0004$, the A element was $P = .065$). This indicates that the monkey may have learned the B element in the Generalization phase rather than the Testing phase.

Finally, in the AxxBx testing phase, the reaction times for both the A and B elements decreased significantly from the first 500 trials of the testing phase trial to the last 500 trials of the testing phase ($p = 0.0001$ and $p = 0.0001$, respectively). Additionally, there was a significant difference the last 500 trials of the generalization phase for both A and B elements when compared to the first 500 trials of the testing phase ($p=.0003$ and $p<.0001$, respectively). This would indicate that she learned the non-adjacent dependency with two intervening elements.

With all three of these sequences being analyzed in the testing and generalization phases, we compared these reaction times to reaction times in both baseline phases. We analyzed whether there was a difference between the reaction times to the specific spatial locations occupied by the A and B elements in the baseline phases compared to the testing and generalization phase(Fig., 8).

First we looked at the reaction time for the A and B element locations in the xxABx sequence and found that there was a statistical difference in the reaction time for only the B element when comparing the testing and generalization phases to both baseline phases (Fig 10., Table 6).

Knowing that learning would have occurred at the end of the testing phase, we used the last 500 trials of the testing and generalization phases and compared these reaction times both

baseline phases (Table 6). The B element in both the testing and generalization phase had a lower reaction time compared to both of the baseline phases. This indicates that the monkey learned to predict the location of the B element, and that she was able to generalize this to novel stimuli. Zoe was able to learn this from recognizing that the A element predicted the B element. She did not learn this by recognizing that the B element was 4th in the sequence because we predict that then would have shown learning for the A element since it was always the 3rd element. This confirms our previous conclusion that Zoe learned the adjacent dependency.

We then ran a statistical test on the A and B positions in the xAxBx sequence to compare both baseline phase reaction times to the testing and generalization phase reaction times and found that there was no statistical difference which confirms that Zoe did not learn the xAxBx non-adjacent dependency (Fig 11., 7 table).

Finally, we ran the same statistical test on the A and B elements in the AxxBx sequence where we compared the A and B reaction times in the testing and generalization phases to the reaction times in the baseline phases and found that there was a statistical decrease in reaction time in the testing and generalization phases for the A element compared to the baseline phases. However, this was not seen in the B element reaction time data(Fig. 12, Table 8). This indicates that Zoe learned the A element from distributional probability because she had recognized that the A element was first in the sequence without previous elements informing her that it would occur. A being at the start of the trial may have aided her learning too.

In conclusion, Zoe was able to learn the adjacent dependency and A element in AxxBx sequence. She was unable to learn any of non-adjacent dependencies(Fig., 9).

Discussion

The results of experiment 2 were based on data from two monkeys and provided mixed results with one learning the adjacent dependency and the other not learning the adjacent dependency.

Cleo, the monkey that failed to learn the adjacent dependency, showed no difference in reaction times for the A and B element in the Testing and Generalization phase compared to the same spatial locations in the baseline phase. This would indicate that she did not learn the adjacent dependency. We predicted that the adjacent dependency would be the easiest dependency to learn, as it requires less working memory demands compared to the non-adjacent dependency (Wilson et al. 2020). With Cleo failing to learn the adjacent dependency, we decided to stop the experiment rather than completing the non-adjacent dependency testing conditions.

Cleo did not learn the adjacent dependency sequence, that is, that the B element always occurred immediately after the A element. Additionally, she was unable to recognize that the A and B elements always appeared 3rd and 4th in the sequence, respectively. It is possible that starting Cleo off on a sequence with fewer elements and building her up to 5 elements may allow her to learn this adjacent dependency, but we found no evidence of learning in this experiment.

The other monkey, Zoe, was able to learn the adjacent dependency, but not the non-adjacent dependencies. We predicted that the adjacent dependency would be easier for the monkeys to learn because less working memory is demanded for these tasks compared to non-adjacent dependency tasks. Zoe was also able to recognize that the A element in the AxxBx sequence was always first. The A element being at the start of the sequence may have aided her in recognizing this, as previous literature has shown that putting dependent elements at the start or end of the sequence can help aid non-adjacent dependency learning (Wilson et al. 2020). However, Zoe was unable to learn that the B element always occurred in the 4th position, or that

it was predicted by the A element. If there were more trials in this phase maybe she would have been able to learn the B element.

In conclusion, one monkey supported our hypothesis that learning dependencies such as the adjacent dependencies are easier to learn than non-adjacent dependencies

General Discussion

The goal of these experiments was to assess the learning of the adjacent and non-adjacent dependencies in rhesus macaques using a serial reaction time paradigm. Given the success of prior SRT paradigms in monkeys demonstrating the learning of at least adjacent dependencies(e.g., Heimbauer et al., 2018), we similarly predicted that our monkeys would learn adjacent dependencies here. Moreover, we predicted that the monkeys may also be able to learn non-adjacent dependencies, although that would likely be more challenging, and less learning might be seen. However, we found mixed evidence across the three monkeys test here.

The results from Experiment 1 initially seemed to suggest that the monkey learned the dependencies between the A and B elements. However, she learned the dependencies immediately in the testing phase and showed equal learning across the dependencies. Both of these results were unexpected, so we further analyzed the results. We found that the faster reaction time to the B element was actually caused by the B elements occurring in a spatial location in which she naturally responded very quickly. When we compared reaction times to the B elements during testing to the same spatial locations during baseline, we found no differences, demonstrating that she did not actually get faster at reacting to these stimuli, and therefore that she did not show evidence of learning the dependencies.

The second experiment was designed to separate each sequence type so that the monkeys only had to focus on one dependency at a time. Additionally, we made the A and B elements blue, rather than red in hopes of aiding the monkeys learning.

In the first monkey, learning of the adjacent dependency did not occur. Adjacent dependencies have been replicable, but here we had a monkey that could not learn it. We analyzed her A and B element reaction times in the Testing and Generalization phase to the same spatial location reaction times in both baseline phases and found there was no difference. This indicated that no learning had occurred. Every monkey is different, and this spatial location task was too difficult for Cleo to learn. If this task was replicated in an audio task, maybe she would show different results.

In the second monkey, she demonstrated learning of the adjacent dependency. We analyzed her reaction times from the start of the testing phase to the end of the testing phase, as we predicted that learning occurs over time. Then we compared her Testing and Generalization phase element B reaction times to the same spatial location in both Baseline phases. In both of those tests, there was a significant decrease in reaction times indicating she learning the B element in the xxABx sequence. We ran the same tests for the A element, but the monkey was unable to show any statistical differences in the Testing and Generalization phases compared to the baseline 2 phase.

Zoe was unable to learn both non-adjacent dependencies. This is because non-adjacent dependencies are more difficult to learn than adjacent dependencies. Non-adjacent dependencies require more working memory demands, and while monkeys have demonstrated that they can learn non-adjacent dependencies, they difficult for them to learn (Wilson et al. 2020).

This monkey did learn the A element in the AxxBx sequence. The A element was the first element in the sequence and previous literature has shown that when dependent elements are on the edge of a sequence (either at the start or end) they can aid the learning of non-adjacent dependencies (Wilson et al. 2020). This ‘edge’ effect (a dependent element at the start or end of a sequence) helped this monkey learn the A element, but not the non-adjacent dependency.

In conclusion with these two experiments, it seems that teaching monkeys adjacent and non-adjacent dependencies is difficult, with only one of three monkeys learning the adjacent dependency and zero monkeys learning non-adjacent dependencies. In the future, it would be interesting to test if reducing the x element spatial locations will allow for better learning. This could be done by reducing the size of the 4x3 grid to a 3x3 or even a different spatial arrangement. Additionally, increasing the amount of trials per sequence will allow the monkeys more time to learn the dependencies.

As more researchers keep experimenting with primates on statistical learning tasks, we will get a better understanding of the cognitive capabilities that these animals have in statistical learning tasks. With this data, we can learn more about the cognitive processes that language utilizes while also finding out what cognitive processes allow for language to arise in humans.

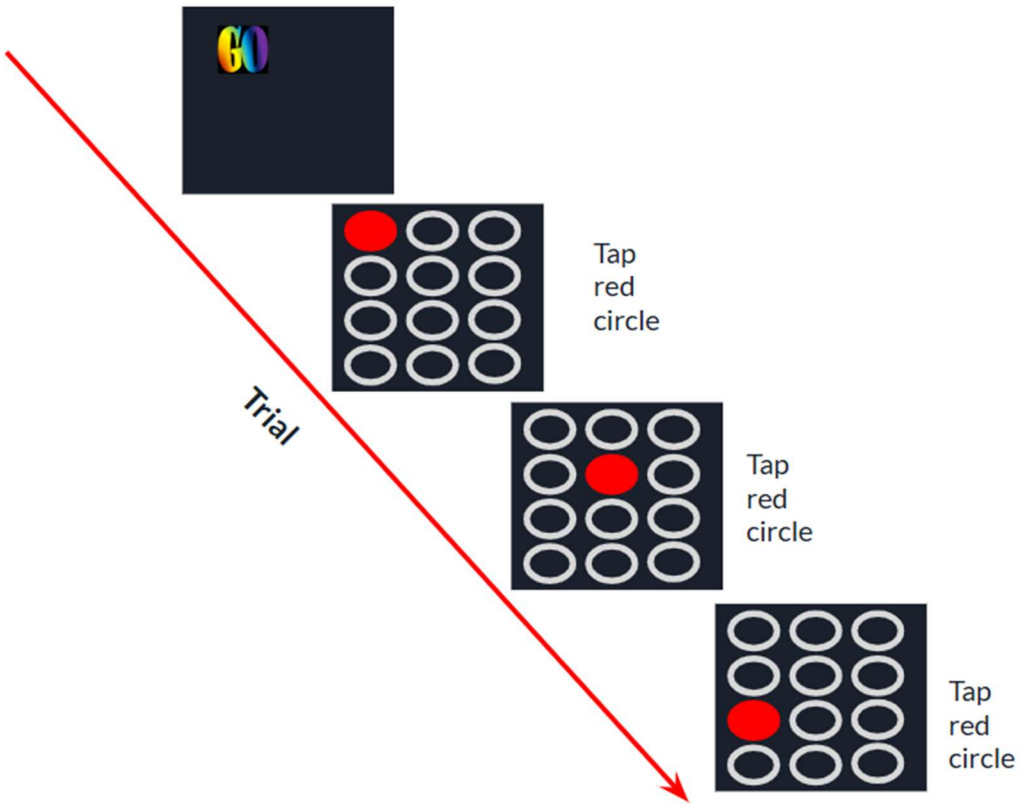

Figure 1: This shows what the task looks like for the monkeys. They will first see the screen on the top left with a 'GO' button and tap that to begin a trial. Then an array of 4x3 dots will appear with one of them being red. The monkey will then hit the red circles until the trial has ended.

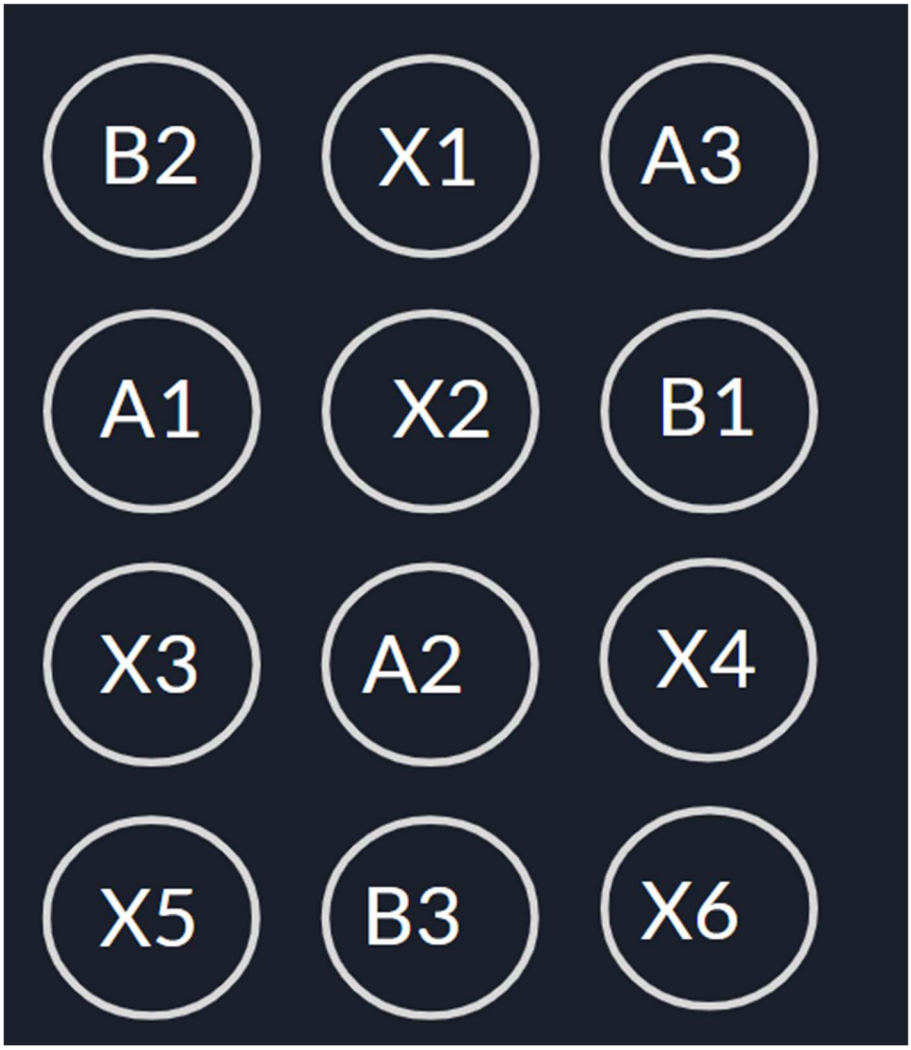

Figure 2: This shows a 4x3 grid with the locations where the stimuli will be. There are 3 A locations (labeled A1, A2, and A3) and 3 B locations (labeled B1, B2, and B3), and 6 X locations (labeled X1-X6). The A and B locations are correlated to a specific sequence and will only occur in 1 of the sequences. For example, A1 and B1 will only appear in the xxABxx sequence. The B1 location will always be predictable because A1 is directly before B1. The Xs will appear in each sequence and will be unpredictable.

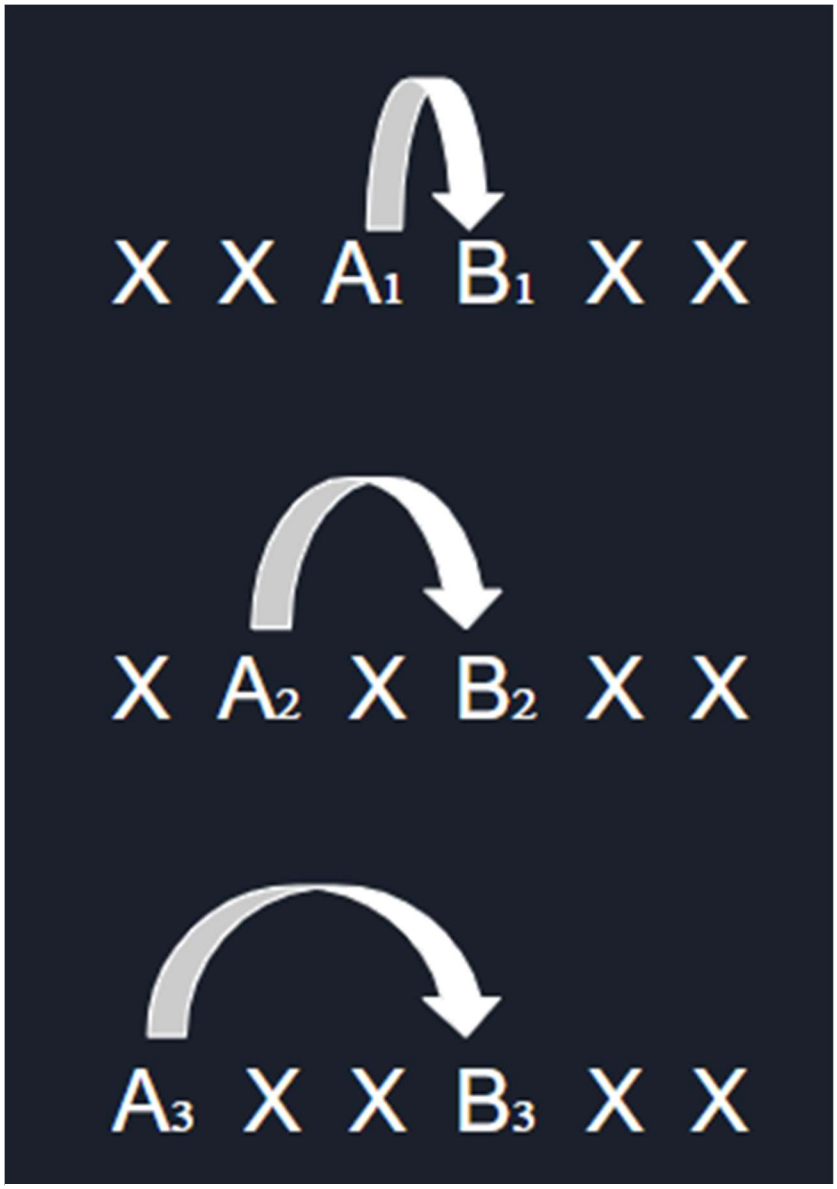

Figure 3: This figure shows the different sequences and when the B element will occur in relation to the A element. In the top sequence, B will occur directly after A, an adjacent dependency In the middle sequence, the B will occur 2 elements after A, a non-adjacent dependency with 1 intervening element. In the bottom sequence, the B element will occur 3 elements after A, a non-adjacent dependency with 2 intervening elements.

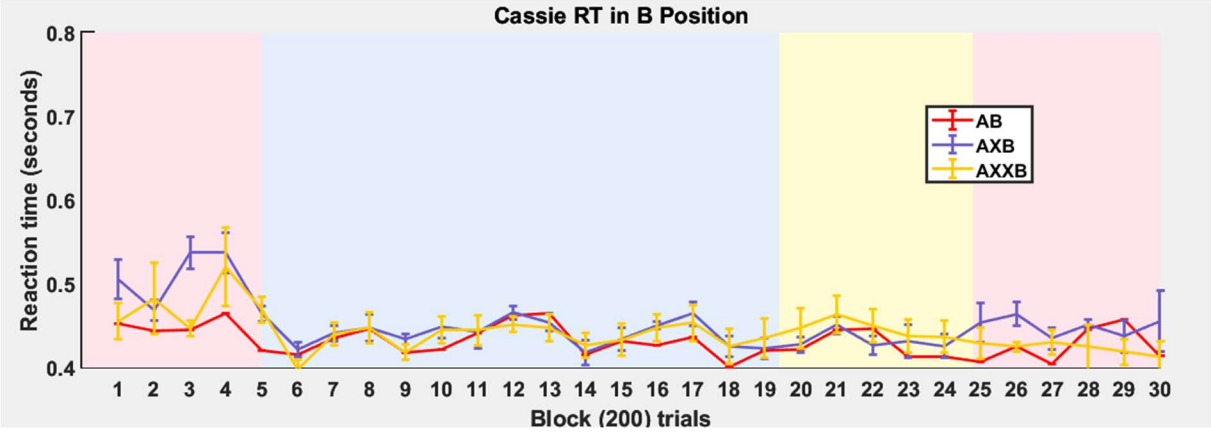

Figure 4: This contains the reaction times for the B element spatial location across every phase for each sequence. The red background denotes the Baseline phases. The blue background represents the Testing phase, and the yellow background represents the Generalization phase.

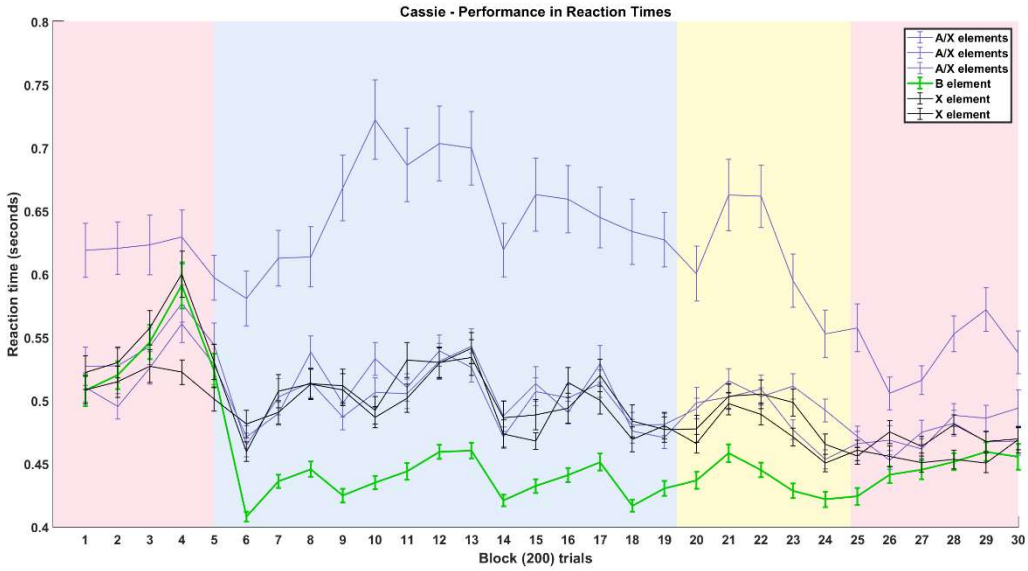

Figure 5: Cassie's performance across phases. Reaction times for each sequence element across each phase are shown. The red backgrounds denote the Baseline phases. The blue background represents the Testing phase. The yellow background represents the Generalization phase. The green line shows the 4th element in the sequence, which corresponds to the predictable 'B' element in the Testing and Generalization phases, and which shows a decrease in reaction time after the Baseline 1 phase.

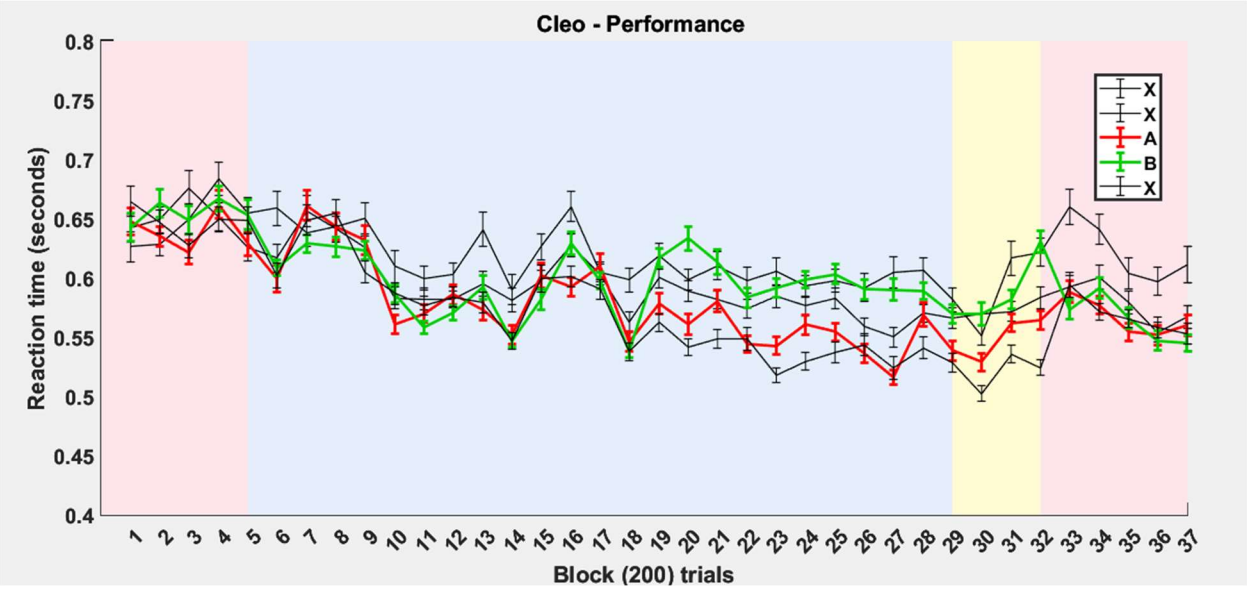

Figure 6: Cleo's performance across phases. Reaction times for each sequence element across each phase are shown. The red background denotes the Baseline phases. The blue background represents the Testing phase, and the yellow background represents the Generalization phase.

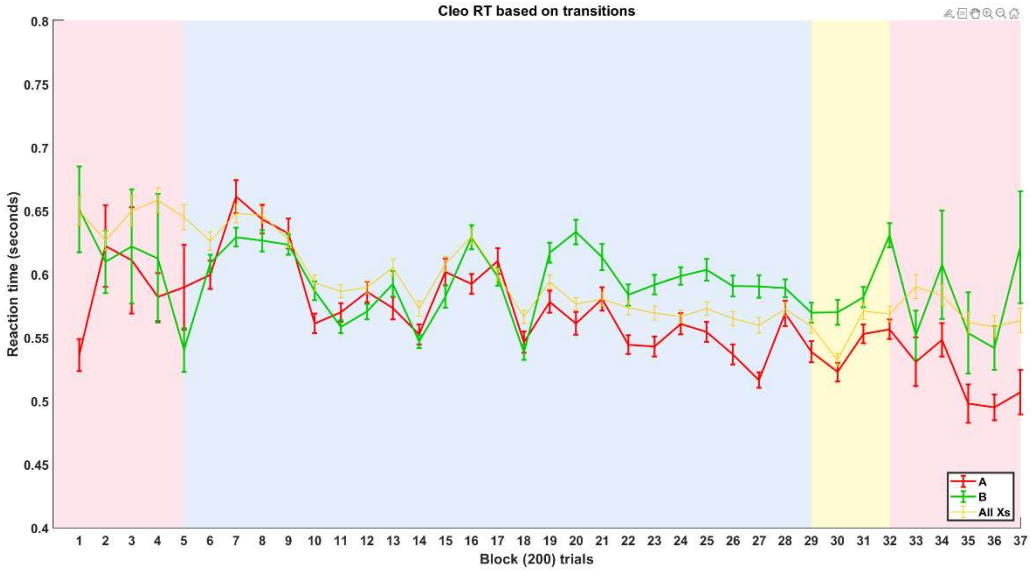

Figure 7: This displays each A and B reaction time across each phase. This includes all of the reaction times to every A and B spatial location. It also includes the X transition average reaction time across each trial block. The red background denotes the Baseline phases. The blue background represents the Testing phase, and the yellow background represents the Generalization phase. To reiterate, this graph is not based on element number in the sequence but based on every reaction time to the A and B element. There is no decrease in reaction time for the B positions, but there is a continual decrease in reaction time for the A position.

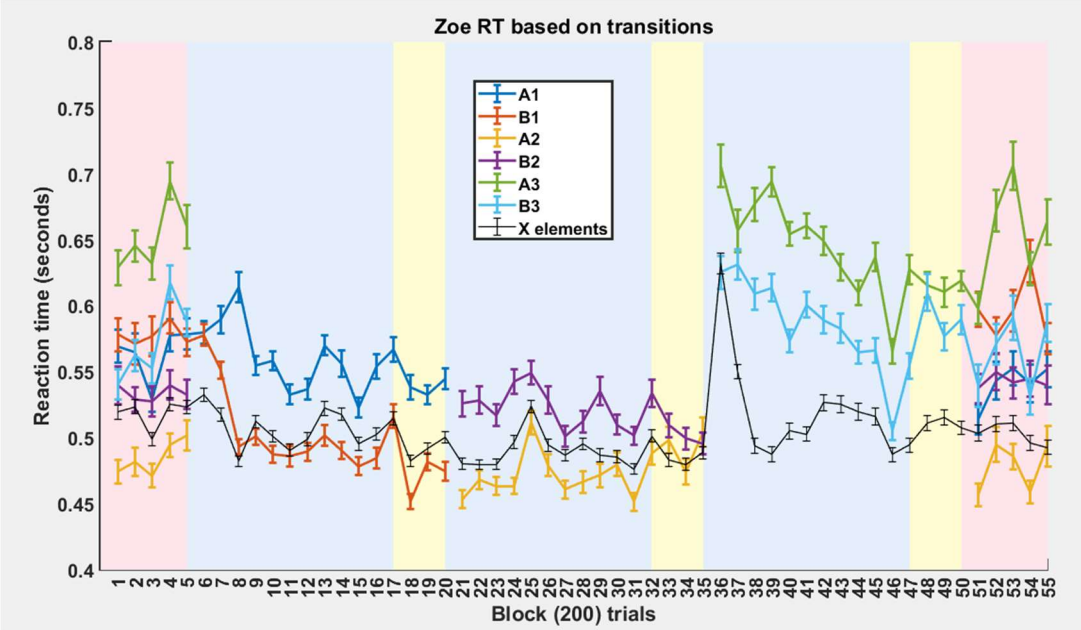

Figure 8: This includes a graph of all of Zoe's reaction times based on their position across the phases. Each A and B element reaction time is shown. The red background denotes the Baseline phases. The blue background represents the Testing phase, and the yellow background represents the Generalization phase. The first sequence (from left to right) is the xxABx sequence, followed by xAxBx sequence, with AxxBx sequence on the right before the Baseline 2 phase.

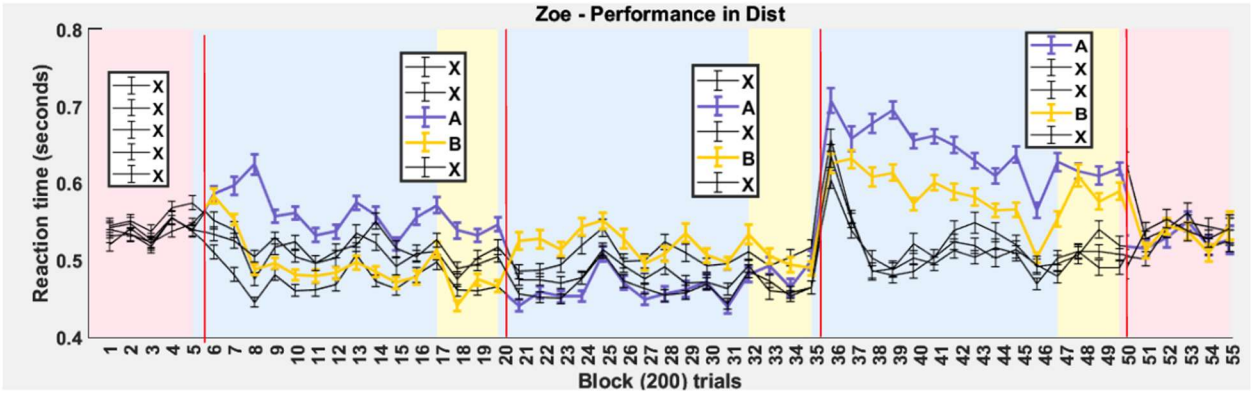

Figure 9: Zoe's performance across phases. The A element is in blue and the B element is in yellow. On the left in red is the baseline phase. Then after the red line, the xxABx Testing phase begins followed by the Generalization phase in yellow. Once the red line appears, the xAxBx Testing phase begins followed by the Generalization phase in yellow. Once the red line appears, the AxxBx Testing phase begins followed by the Generalization phase in yellow. Once the red line appears, the Baseline 2 phase occurs. The legends are for the elements within the red lines and the A element changes ordinal elements in the sequence, along with A and B changing positions on the 4x3 grid.

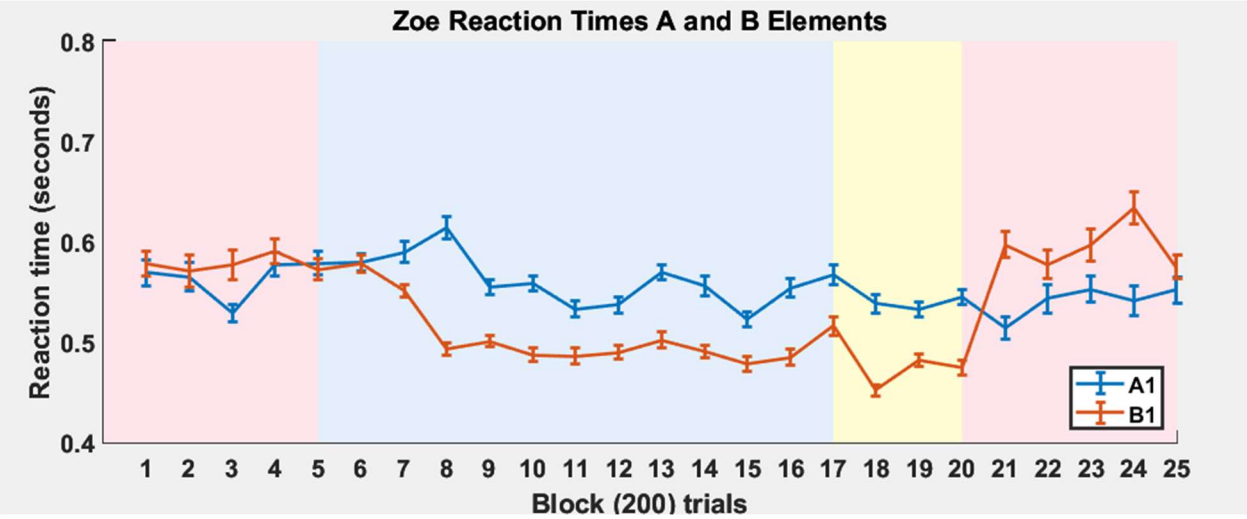

Figure 10: The reaction times for the A and B elements spatial location in the xxABx sequence type is plotted. Learning can be seen in the B element spatial location because of a decrease in reaction time across the testing and generalization phase compared to both baseline phases. There is no learning occurring with the A element spatial location.

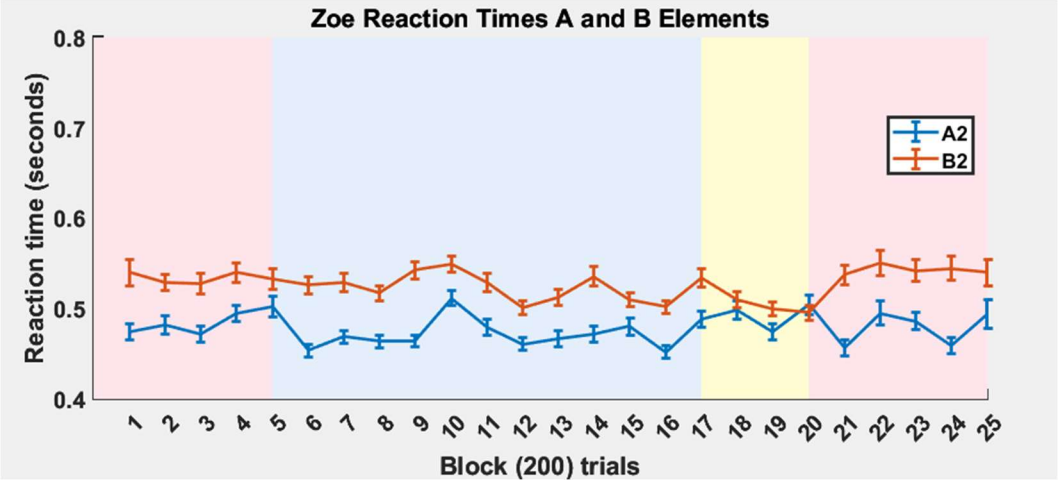

Figure 11: The reaction times for the A and B elements spatial location in the xAxBx sequence type is plotted. There is no decrease in reaction time in either the A and B element spatial location, which indicates no learning has occurred.

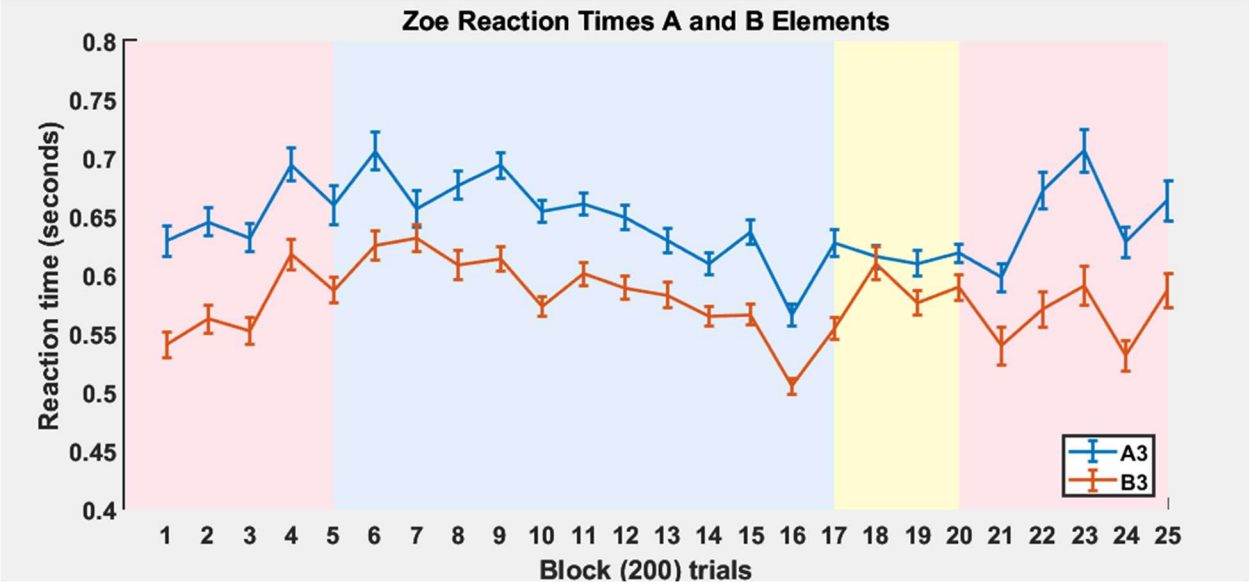

Figure 12: The reaction times for the A and B elements spatial location in the AxxBx sequence type is plotted. There is learning seen in the A element, as there is a decrease in reaction time from the testing and generalization phase compared to both baseline phases. There is no significant decrease in reaction time in the B element compared to both baseline phases, indicating no learning has occurred.

Table 1: Results of permutation tests in each phase of the experiment, which compares each element reaction time to element 4 reaction time. In the Baseline phases, none of the stimulus elements are predictable based on other elements of the sequence, so we would predict no differences. We used position 4 as the element for comparison, because in the subsequent Testing and Generalization phases, the predictable B element always occurs in position 4, so in those phases we expect the fastest responses to stimuli in this sequence position. Significant differences are shown in bold.

Element comparison	Position 1 and 4	Position 2 and 4	Position 3 and 4	Position 5 and 4	Position 6 and 4
Baseline 1	**P<0.0001**	P=0.2574	P=0.9428	P=0.9979	P=0.1481
Testing	**P<0.0001**	**P<0.0001**	**P<0.0001**	**P<0.0001**	**P<0.0001**
Generalization	**P<0.0001**	**P<0.0001**	**P<0.0001**	**P<0.0001**	**P<0.0001**
Baseline 2	**P<0.0001**	**P<0.0001**	**P<0.0001**	P= 0.0752	**P<0.0001**

Table 2: This displays the P values for a permutation test done across the different sequences for the B element reaction time. We applied Bonferroni corrections for multiple comparisons. None of the B elements reaction times are statistically different.

Sequence comparison of B element	xxABxx vs. xAxBxx	AxxBxx vs. xAxBxx	AxxBxx vs. xxABxx
P values	P = 0.1041	P =0.2340	P = 0.9890

Table 3: This contains the P values for the permutation test that compares the B element reaction times across 2 phases per unique sequence. We applied Bonferroni corrections for multiple comparisons. The bold data is significant. This data table shows 4 statistically significant results. 2 of them occur when comparing the Baseline 1 B element reaction times to the Testing phase B element reaction times. This significant decrease in reaction time may be occurring because Cassie is just getting faster at the SRT task. The other statistically significant result occurs when comparing the Baseline 2 phase reaction times to the generalization phase reaction times for the B element and this is because in the Baseline 2 phase, the B element reaction time continues to decrease, indicating that this is not because of learning but because she is just getting faster at the SRT task. The other statistically significant result is in the Baseline 1 and Generalization phase B element reaction time comparison which shows the B element decreasing in the Generalization phase, but this phase is not statistically different from Baseline 2 phase B element, so no learning has occurred.

Comparisons	xxABxx B element	xAxBxx B element	AxxBxx B element
Baseline 1 vs. Testing	P= 1	p **< 0 .0001**	**p < 0 .0001**
Testing vs Baseline 2	P= 1	P = 0.2265	P = 0.2745
Baseline 1 vs Generalization	P= 1	**p < 0 .0001**	P= 0.3315
Baseline 2 vs Generalization	P= 1	P= 0.1395	**P= 0.0195**
Testing vs Generalization	P= 1	P= 1	P= 0.5085

Table 4: This includes a permutation test of all of the Baseline 1 phase reaction times, comparing each element to another one. We applied Bonferroni corrections for multiple comparisons. None of the elements have a statistically different reaction time in the Baseline phase.

	Element 1	Element 2	Element 3	Element 4
Element 1				
Element 2	P = 4.532			
Element 3	P = 1.106	P = 0.91		
Element 4	P =1.91	P = 2.00	P = 0.11	
Element 5	P = 4.53	P = 4.22	P = 1.04	P = 1.43

Table 5: Displayed are the P values for a permutation test for the A and B element across reaction times across different phases. Bolded are the significant results. In comparing element A Baseline 2 reaction times to Testing phase reaction times, the Baseline 2 phase has faster reaction times than the Testing phase reaction times so this result does not indicate learning, but just getting faster at the task. In the Baseline 1 phase to Generalization phase comparison in the A element, there is a decrease in reaction time in the Generalization phase indicating that learning may be occurring, but when comparing Generalization phase to Baseline 2 phase reaction times, the A element reaction time gets faster in the Baseline 2 phase so no learning is occurring. There is a decrease in reaction time for the A element from the Testing phase to Generalization phase, but this does not indicate learning because of the decrease in reaction time that continues in Baseline 2. The B element reaction time shows no statistical differences between any phases of the experiment.

Comparison between phases	Baseline 1 vs testing	Baseline 2 vs testing	Baseline 1 vs generalization	Baseline 2 vs generalization	Testing vs generalization
A element	P = 0.1151	**P > 0.0001**	**P > 0.0001**	**P > 0.0001**	**P >0.0001**
B element	P = 0.2239	P = 0.1697	P = 0.2384	P = 0.2183	P = 0.4752

Table 6: This includes the P values for a permutation test for the A and B elements across each phase of the experiment in the xxABx sequence. Bold data is significant. For the A element, there was a decrease in reaction time in the Baseline 2 phase compared to the Testing phase, indicating Zoe getting better at the SRT experiment. There is also a decrease in reaction time for the A element from Baseline 1 to Generalization phase, but this decrease is not significant from Baseline 2 phase, indicating no learning. For element B, there was a significant decrease in reaction time between the testing and generalization phases and both baseline phases, indicating learning has occurred.

xxABx sequence elements	Baseline 1 vs. Testing	Baseline 2 vs. Testing	Baseline 1 vs. Generalization	Baseline 2 vs. Generalization
A element	P = 0.4802	**P =0.0129**	**P = 0.0029**	P = .3889
B element	**P = 0.0041**	**P = 0.0278**	**P = 0.0028**	**P = 0.0002**

Table 7: This includes the P values for a permutation test for the A and B elements across each phase of the experiment in the xAxBx sequence. For the A element, there no statistically significant change in reaction time across the phases. This is true too with the B element reaction times. These results indicate that there was no learning occurring for the A and B elements in the xAxBx sequence.

xAxBx elements	Baseline 1 vs. Testing	Baseline 2 vs. Testing	Baseline 1 vs. Generalization	Baseline 2 vs. Generalization
A element	P =.1453	P = .4555	P = .2707	P = .0751
B element	P = .2439	P = .0873	P = .1944	P = .4072

Table 8: This includes the P values for a permutation test for the A and B elements across each phase of the experiment in the AxxBx sequence. Bolded results are statistically significant For the A element, there was a statistical decrease in reaction time in both the Testing and Generalization phase compared to both Baseline phases. This indicates that Zoe learned the A element. For the B element, there was a statistical difference from Testing to Baseline 1 phase, but there was no difference in Testing to Baseline 2 phase reaction times, so no learning occurred in the testing phase. Similarly in the Generalization phase, there was a significant decrease in the B reaction time compared to Baseline 1, but when compared to Baseline 2, the Baseline 2 phase reaction times were statistically faster than the Generalization phase which indicates no learning.

AxxBx elements	Baseline 1 vs. Testing	Baseline 2 vs. Testing	Baseline 1 vs. Generalization	Baseline 2 vs. Generalization
A Element	**P = .0001**	**P = .0098**	**P = .0002**	**P = 0.0466**
B Element	**P = .0002**	P = .3587	**P = 0.0369**	**P = .0403**

Works Cited

Fitch, W. Tecumseh, and Marc D. Hauser. “Computational constraints on syntactic processing in a nonhuman primate.” *Science*, vol. 303, no. 5656, 16 Jan. 2004, pp. 377–380, https://doi.org/10.1126/science.1089401.

Gebhart, Andrea L., et al. “Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds.” *Psychonomic Bulletin & Review*, vol. 16, no. 3, June 2009, pp. 486–490, https://doi.org/10.3758/pbr.16.3.486.

Gomez, Rebecca L, and LouAnn Gerken. “Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge.” *Cognition*, vol. 70, no. 2, Mar. 1999, pp. 109–135, https://doi.org/10.1016/s0010-0277(99)00003-7.

Heimbauer, Lisa A., et al. “Visual artificial grammar learning by Rhesus Macaques (macaca mulatta): Exploring the role of grammar complexity and sequence length.” *Animal Cognition*, vol. 21, no. 2, 12 Feb. 2018, pp. 267–284, https://doi.org/10.1007/s10071-018-1164-4.

Malassis, Raphaëlle, et al. “Non-adjacent dependencies processing in human and non-human primates.” *Cognitive Science*, vol. 42, no. 5, 20 May 2018, pp. 1677–1699, https://doi.org/10.1111/cogs.12617.

Nissen, Mary Jo, and Peter Bullemer. “Attentional requirements of learning: Evidence from performance measures.” *Cognitive Psychology*, vol. 19, no. 1, Jan. 1987, pp. 1–32, https://doi.org/10.1016/0010-0285(87)90002-8.

Pacton, Sébastien, et al. “Is an attention-based associative account of adjacent and nonadjacent dependency learning valid?” *Acta Psychologica*, vol. 157, May 2015, pp. 195–199, https://doi.org/10.1016/j.actpsy.2015.03.002.

Reber, Arthur S. “Implicit learning of artificial grammars.” *Journal of Verbal Learning and Verbal Behavior*, vol. 6, no. 6, Dec. 1967, pp. 855–863, https://doi.org/10.1016/s0022-5371(67)80149-x.

Saffran, Jenny R., et al. “Statistical learning by 8-month-old infants.” *Science*, vol. 274, no. 5294, 13 Dec. 1996, pp. 1926–1928, https://doi.org/10.1126/science.274.5294.1926.

Santolin, Chiara, and Jenny R. Saffran. “Constraints on statistical learning across species.” *Trends in Cognitive Sciences*, vol. 22, no. 1, Jan. 2018, pp. 52–63, https://doi.org/10.1016/j.tics.2017.10.003.

Versace, Elisabetta, et al. “Positional encoding in cotton-top tamarins (Saguinus Oedipus).” *Animal Cognition*, vol. 22, no. 5, 1 July 2019, pp. 825–838, https://doi.org/10.1007/s10071-019-01277-y.

Wilson, Benjamin, Michelle Spierings, et al. “Non-adjacent dependency learning in humans and other animals.” *Topics in Cognitive Science*, vol. 12, no. 3, 8 Sept. 2018, pp. 843–858, https://doi.org/10.1111/tops.12381.

Wilson, Benjamin, William D. Marslen-Wilson, et al. “Conserved sequence processing in primate frontal cortex.” *Trends in Neurosciences*, vol. 40, no. 2, Feb. 2017, pp. 72–82, https://doi.org/10.1016/j.tins.2016.11.004.

Wilson, Benjamin, Yukiko Kikuchi, et al. “Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans.” *Nature Communications*, vol. 6, no. 1, 17 Nov. 2015, https://doi.org/10.1038/ncomms9901.