The Role of Context in Object-Word Learning: Findings from Computational Modeling

Bachelor's Project Thesis

Tinke van Buijtene, s2701421, t.van.buijtene@student.rug.nl

Supervisors: Dr. Jacolien van Rij & Dr. Jennifer Spenader

Abstract: It is yet unknown how exactly children are able to match objects they perceive with their correct labels. This is known as the mapping problem. Researchers have proposed various theories on how this problem could possibly be dealt with, two of the most well-known being cross-situational learning and propose-but-verify. But one aspect is often overlooked: the context in learning environments. For this study, I have built computational models of cross-situational learning and propose-but-verify that take on experiments done by Dautriche and Chemla (2014). Using these models gave me the opportunity to inspect closely how different contextual setups in the experiments affect the performance of both strategies. The models' performances were then compared to human data. Results showed that even with the context modulations, neither necessarily resembles human data better than the other. I discuss one possible explanation for this finding: that people actually use a combination of both strategies.

Keywords: object-word learning, cross-situational learning, propose-but-verify, computational modeling

1. Introduction

Young children learn words in their native language very rapidly. Upon becoming 2 years of age, children reach an average vocabulary of 300 words (Fenson et al., 2014). This is an impressive achievement, because the environments in which children learn words are often rather complex, consisting of many different objects, concepts and ongoing actions. It is still unknown how exactly children match the words they hear with what they perceive in the world. This issue is also known as the mapping problem (Bloom, 2000). Quine (1960) similarly states that for each naming event, there is referential ambiguity. A classic example that he has given to illustrate this phenomenon is that of a native speaker of a foreign language, who upon seeing a rabbit utters the word *"gavagai"*. To someone unfamiliar with this language, the word may refer to the rabbit, but it may as well refer to "animal", "food", "grass" or any other present object. All the more, it may even refer to something that is not present at the moment, such as "hunting". Quine's aim was to show that, in theory, there might be an infinite set of possible candidate meanings for an unknown word. However, in this example we can already sense that it is more likely that *gavagai* means "rabbit", rather than "hunting". It is plausible to think that learners exclude rather spurious candidate meanings from the number of considered possibilities, which would greatly reduce the ambiguity. Though in spite of this reduction, a reasonable amount of possibilities ceases to exist. The number of different possibilities results in uncertainty. The learner, due to this uncertainty, requires a means to distinguish the true referent of a word from the other possibilities, also known as distractors. The means to overcome referential uncertainty could be realized as an underlying mechanism or a strategy.

In previous studies, researchers have come up with several explanations on how learners cope with the referential ambiguity in word-learning. These each have their differences, but can be classified into two general types of accounts.

On one hand, several researchers have proposed that learners are able to keep track of statistics of word-object co-occurrences across diverging learning environments or situations. In other words, according to this theory learners can in some way carry information from previously encountered learning situations to another. For this fact, one of the more prevalent theories in this domain is called ***cross-situational learning***.

In theory, when a new word is uttered, there is a set of candidate meanings of that word due to the referential uncertainty. But when the word is repeatedly presented in new situations, the different contexts of those situations can be used to rule out candidate meanings which will reduce the uncertainty (Smith & Smith, 2012). This already suggests that context plays an important role in object-word learning.

On the other hand, there are researchers that propose that a learner does not keep track of information from previously encountered learning situations. The theory that has been proposed by Trueswell et al. in 2013 is called ***propose-but-verify***, in which learners upon hearing a new word, propose a hypothesis, which may be verified upon later encounters of that word. If their hypothesis is verified, i.e. the hypothesized object-word match is shown on a new trial, the hypothesis is strengthened. If, however, the word is heard in absence of the hypothesized object, the hypothesis is rejected. In this case a new hypothesis is chosen. We may ask ourselves how a learner chooses a (new) hypothesis and whether the context of a learning situation has any influence on the process.

My aim in this study was to investigate the role of context in object-word learning, using the two theories proposed above. I constructed computational models of the two theories and used a previously done study by Dautriche and Chemla (2014) to assess both of the models' performances. The named study that Dautriche and Chemla carried out, consisted of different modifications of the same experiment in which the context differed. Their study was interesting to use for us, because this helped us to obtain the models' performances across different contextual setups. The performances of our computational models were tested in the same experiments and compared to the human data.

Having obtained the results, we made two types of comparisons: a comparison between the two models (model-to-model comparison) and a comparison between each of the models and the human data that was obtained from Chemla and Dautriche's experiments (model-to-data comparison). We found that the outcome of the two models is mostly similar with respect to the choices they make during the experiment, with minor differences. When comparing the models to the human data, we find that neither necessarily resembles human data better than the other. I discuss one possible explanation: that learners in reality use a strategy or mechanism that is in fact a combination of cross-situational learning and propose-but-verify.

2. Background

It is worth to mention that the two theories of object-word learning that we will be comparing in this experiment, cross-situational word learning and propose-but-verify, can be classified more broadly as two different learning accounts. Cross-situational word learning can be seen as a form of implicit learning, whereas the propose-but-verify strategy uses propositional logic to reduce uncertainty.

In cross-situational word learning, learners are eventually able to determine the correct object-word pair by choosing the object that has the highest occurrence with the referring word as, implicitly, that word will be the most associated with the object in comparison to the other words. Figure 1 illustrates a simplified example of how uncertainty can be reduced over multiple exposures. On each exposure, i.e. each time the target word "horse" is used, there is a set of objects present that become candidate meanings. The appearance of the real horse is consistent over multiple exposures and therefore this candidate will be associated more with the target words than the other candidates.

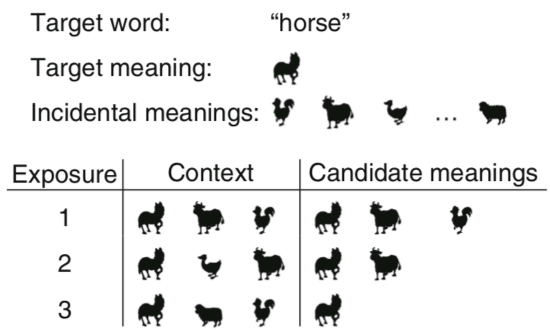

Figure 1: Simplified example of how uncertainty is reduced in cross-situational word learning

The implicit learning factor of the cross-situational learning theory may be realized by using, for example, the Rescorla-Wagner model, a theory of classical conditioning (Rescorla & Wagner, 1972). Rescorla-Wagner is a variation on Pavlovian classical conditioning, where we do not assume that conditioned stimuli followed by unconditioned stimuli result in learning.

This is based on findings by Kamin (1969) on the blocking effect. Namely, that once one stimulus has a strong association with a certain outcome, any other stimulus that is presented at the same time will not be learned to be a predictor for the outcome. In the Rescorla-Wagner model, the change in associative strength of a stimulus depends on the existing associative strength of that stimulus and of the other present stimuli. If a stimuli is already strongly associated with a certain response, the association will not strengthen as much as when the stimuli is not yet strongly associated with the response. It is important to emphasize that the Rescorla-Wagner model that we use does not merely implement associative learning, but rather a form of discriminative learning. The discriminative part here implies that if a stimulus is not present, there is a process of unlearning. These concepts of the theory are realized through the mathematical formula in equation 1, which Miller et al. (1995) introduced in an assessment of the Rescorla-Wagner model.

$$\Delta V_X^{n+1} = \alpha_X \beta(\lambda - V_{tot}) \quad (1)$$

In the equation above, ΔVx resembles the change in associative strength between the conditioned stimulus (labeled *X*) and an unconditioned stimulus. The change occurs from the current trial to the next, hence the change can be seen on trial *n+1* (where *n* resembles the number for the current trial). The change is given by the associability of the conditioned stimulus (α), bounded between 0 and 1, multiplied by the associability of the unconditioned stimulus (β), bounded between 0 and 1. Typically, α and β have a value of 0.1. The associability of a stimulus is closely related to the intensity of that stimulus. The λ resembles the maximum associative strength that the unconditioned stimulus can support. Because we are using discriminative learning, which makes it possible for weights to become negative, the value for λ is 0 when an outcome is not present. V_{tot} yields the total sum of associative strengths of all conditioned stimuli (including stimulus X) that are present on trial *n+1* (Miller et al., 1995).

Using the Rescorla-Wagner model, the mechanism of cross-situational learning can be realized as a linked network consisting of associative information. As such, the theory can be seen as an implicit learning account.

There are opponents of cross-situational learning, who argue that a flaw in this theory is that it makes the assumption that children can somehow keep track of multiple hypotheses for a word, which would require a vast amount of memory (see Medina et al., 2011; Trueswell et al., 2013). Instead, they propose, that a learner does not keep track of information from previously encountered learning situations and thus that the type of learning is not associative, they propose that the learners in fact use propositional logic to learn object-word pairings. The theory that has been proposed by Trueswell et al. in 2013 is called ***propose-but-verify***, in which, as its name suggests, learners per learning situation propose a hypothesis which may be verified upon later learning situations. Using this strategy, learners only have to keep track of their current hypothesis, which is a single object-word match. The hypothesis can either be verified, which happens when an object-word match is shown on a new trial, in which case the hypothesis is strengthened. If, however, the target word is heard in absence of the hypothesized object, the hypothesis is rejected. In that case a new hypothesis is chosen.

Trueswell et al. (2013) describe the specific steps that are taken in the strategy as follows:

1. Begin by guessing at chance.
2. On an additional occurrence of a word, remember the previous guess with some probability α.
3. If the remembered guess is present in the current referent set (i.e., confirmed), increase

the value for α and select the referent; otherwise select a referent at random.

This strategy thus knows only one free parameter, α, that resembles that probability that a learner's previous guess of a target word is remembered.

We have just described two accounts of how learners could cope with the referential uncertainty (i.e. the mapping problem) that is found in object-word learning. On one hand we have the associative account, or cross-situational learning, which proposes that learners compare conjectures that they gather over multiple situations. The core difference with the propositional logic account, propose-but-verify, is that according to the latter theory learners do not keep track of information about previously encountered situations - instead they store only a single conjecture, being their hypothesis. This hypothesis is carried forward until disproven by later encounters.

While research has been focused on the possible underlying mechanisms of object-word learning, one source of information is often left out: the ***context*** of the those learning situations. A previously conducted study by Horst (2013) has demonstrated that children's word learning benefits from contextual repetition. Horst investigated the influence of contextual repetition on word-learning by reading groups of children either the same three storybooks repeatedly or nine different storybooks once each. They found that all children performed well on the tests they made right after the storybook reading, however only the group of children who had listened to the same storybooks repeatedly retained those word-object associations upon a second test, one week after the storybook readings, meaning that contextual repetition proved beneficial for learning to occur. Horst states that this may also explain why stories or songs that make use of repetitive features are more appealing to children.

In this study, we want to further explore the influence of context on object-word learning. We are specifically interested in the effect of context modulation on the learning rate of words using the cross-situational learning and propose-but-verify mechanisms.

We have chosen to use computational modeling as a tool to investigate the role of context in object-word learning. In prior studies, several researchers have laid out experiments with human participants in which the context was modulated in some way to test its effects on object-word learning (Chemla & Dautriche, 2014; Roembke & McMurray, 2016 - see section "Previous studies with human subjects"). With our modeling approach, we aim to provide new insights to current findings.

Computational modeling has two main advantages as opposed to testing with human subjects. Firstly, modeling allows us to control the learning strategies that we want to investigate. We have the possibility to see the outcome of a strategy at each stage during the learning process. With human participants, this would not be possible, as we would only be able to see the outcome of the tests because of which we can only take a guess at what sort of strategy would have produced that outcome. The other advantage is that with the aid of modeling, we can examine how learning progresses over a longer period of time or with more exposures. This is difficult to do with human subjects, whose time and attention is limited.

Previous studies with human subjects

Several studies have conducted language experiments with the aim to gain more insight into the underlying processes of object-word learning (see e.g. Smith et al., 2010; Ramscar et al., 2013; Trueswell et al., 2013). In these experiments, the participants (children or adults) were typically shown different objects along with a to-be-learned word. Analyzing what choices participants make when selecting what they believe is the referent of a word, allows us to infer what underlying learning mechanism or strategy they might have been using.

We found two studies that were not just focused on the mechanism behind object-word learning, but rather on how the learning was effected by different contexts.

In 2014, Dautriche and Chemla conducted a study in which their aim was to investigate how context could modulate learners' strategies during object-word learning and how that influences the strategies' efficiencies. They set out three experiments in which they had participants look at pictures upon hearing a word, with the aim to have the participants eventually link that word to the corresponding object. They would show participants four pictures at a time per trial, where the correct object would always be present in learning instances for that word. This was the basic set up. The difference between the three experiments was the context, which they tried to simulate by altering which objects were shown together. In their first experiment, the objects that were shown together at a time were entirely randomized. In the second experiment, they chose to only show pictures that belonged to a same category (e.g. animals or clothes) together, so as to simulate a natural context (e.g. a zoo or clothing store context). In the third experiment, Dautriche and Chemla wanted to further explore different kinds of contexts and their effects by grouping objects into consistent contexts (consisting of objects that have no semantic coherence) instead of natural contexts. From this study, they found that learning benefits from situations that consistently contain members from a certain group, both when this group represents a natural context and when the context is merely consistent without any semantic coherence. These results suggest that a consistent context serves as a memory cue during the learning process.

Roembke and McMurray conducted a follow-up study in 2016 to look further into this effect. Their aim in this study was to simulate contexts in a more natural, probabilistic manner by not having the context appear on only the first block, like Dautriche and Chemla chose to do. The reason that Roembke and McMurray chose for a different approach on simulating context was that they believed that Dautriche and Chemla's approach made context very salient. Roembke and McMurray therefore tried to manipulate contextual consistency between trials in a more natural or probabilistic manner. With this they investigated how contextual consistency could not only serve as an aid to learning, but also as a cost.

We want to build on the findings by other studies such as those by Dautriche & Chemla, and Roembke & McMurray. We are interested in what computational modeling can tell us about the role of context in object-word learning using cross-situational learning and propose-but-verify.

We want to evaluate our results by comparing them to data from human participants. This is why we chose to model a previously carried out experiment on context in object-word learning. Recreating the setup from an actual experiment gives us an advantage as this makes it easy to compare our models' results to the results of that same experiment done with human participants. As mentioned before, there have been two research groups that tried to investigate the role of context in a similar way to our current study. The studies that were conducted by Dautriche and Chemla on the one hand, and Roembke and McMurray on the other hand, were both suitable for our modeling project. However, for practical reasons, we favored modeling the study done by Chemla & Dautriche. The experiments that Roembke & McMurray carried out involved a lot of trials, which would be more complex to capture in a model because the parameters that we are using might change over time.

We have chosen to model only the first two of the experiments that Chemla & Dautriche carried out. The reason for this is that comparing the results of these two experiments are sufficient to gain insight how the context influences the learning. The first two experiments that Dautriche and Chemla carried out are crucial for our investigation, as they show the differences between learning in a situation where context plays no role and learning in a situation where context occurs and therefore may serve as a memory cue.

3. Method

For this study I have chosen to model both experiments using the R software (R Core Team, 2017). I have created a package containing the necessary functions to represent the experiments and the learning mechanisms of cross-situational learning and propose-but-verify.

We have modeled two different modifications of the same experiment that Dautriche and Chemla carried out (see sections 4 and 5), and since they were so alike we were able to use the same functions to simulate both experiments. I made use of the packages ndl (Arppe, A., Hendrix, P., Milin, P., Baayen, R. H., Sering, T. & Shaoul, C., 2015) and NDLvisualization (van Rij, 2018) to simulate the mechanism of cross-situational learning. Since such a package did not exist yet for propose-but-verify, I have written several R functions that were necessary to simulate the strategy. I have based these functions on the strategy as how it was described by Trueswell et al. (2013) (see Background).

3.1 Cross-situational learning

The cross-situational learning model we have constructed learns based on the cues that are present in a learning event. In all learning events, multiple cues are presented together with an outcome. The outcome is the to-be-learned word and the cues are the presented objects (pictures), alongside a background cue, which we added manually to each learning event. Furthermore, since category plays a role in one of our experiments, we also added cues for the categories of the objects. The presented cues are mapped to the outcome and a list of all the mappings in an experiment forms the data frame that is used for the learning. This data frame is fed to a function RWlearning (ndl and NDLvisualization packages) that implements the Rescorla-Wagner model, which is a form of the mechanism behind cross-situational learning. In this implementation of the Rescorla-Wagner model, we use values of 0.1 for both the associability of the conditioned stimulus (α) and the associability of the unconditioned stimulus (β). The function returns a list with a weight matrix for each learning event. The weight matrices represent the strengths of association of the network, which are all cue-outcome combinations. The difference in weights over time allows us to see the learning progress.

3.2 Propose-but-verify

In modeling the mechanism of the propose-but-verify theory, we followed the steps that Trueswell et al. (2013) describe (see section 2). The first of these steps states that a choice is made at random upon the first trial, as no previous information is known. One of the presented objects during that event is chosen and the object-word match is kept as a hypothesis. Upon a new learning situation for the same word, we implemented two consequent options:
First, there exists a chance that the object chosen on the last trial is remembered. This chance is expressed by the probability of α, which Trueswell et al. (2013) proposed to have a value of 0.26 initially. They proposed this value based the percentage of participants that had been correct on the preceding trial, but not on the current trial, which gave them an indication of the probability that participants remember their hypothesis. If the object that was chosen on the last trial is remembered and currently present, it is chosen again. If this is the case, the value for α increases and becomes 0.71, which is also based on findings by Trueswell et al. (2013). This is another value they found from the percentage of participants that after several trials were correct on the preceding trial, but not on the current one.
If the object is remembered but not present, or if it is not remembered at all, we implemented a second probability. With a probability of α again, there is a chance that the semantic category that the previously chosen object belonged to is remembered. If the category is remembered and one object from that category is present on the current trial, it is chosen. If multiple objects from that category are present, one of those objects is chosen at random. If neither the object nor the category is remembered, or if they are but no object belonging to the same category is present in the current set, a new object is chosen at random. Figure 2 shows a visual representation of the steps taken by our propose-but-verify model.

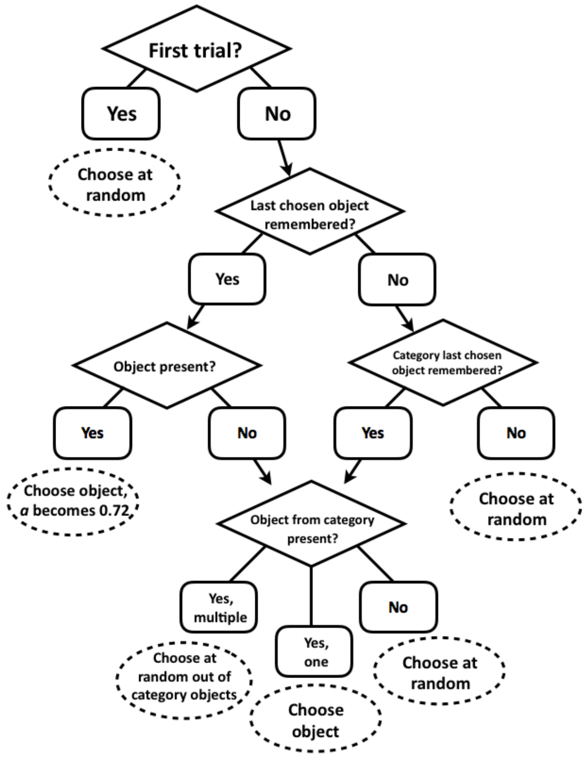

Figure 2: Decision tree that entails the steps taken by the propose-but-verify model

4. Experiment 1

4.1 Introduction

The first experiment that we carried out was a computational model of Dautriche and Chemla's (2014) Experiment 1. In their Experiment 1, they aimed to represent a learning situation in which various concepts are presented to the subject. They did this by showing pictures over different trials, in which subjects saw a word at the same time that referred to one of the shown pictures. The setup of the experiment was similar to the study done by Trueswell et al. in 2013. Trueswell et al. used a paradigm in which subjects heard the to-be-learned words while being shown a five pictures of different objects. Dautriche & Chemla reduced this to a number of four pictures per word. With this experiment, Dautriche & Chemla wanted to show how object word-learning progresses over trials, without taking the influence context into account. In the computational model of this experiment that we made, we used the exact same setup as Dautriche & Chemla.

4.2 Design

There were a total of twelve to-be-learned words. These words were legal English non-words and they represented 12 different objects. The non-words were *blicket, dax, smirk, zorg, leep, moop, tupa, krad, slique, vash, gaddle* and *clup*. They represented the following 12 objects correspondingly: *cat, dog, cow, rabbit, pants, hat, socks, shirt, pan, knife, bowl,* and *glass*. The objects were shown in trials, in which a target object would be presented along with 3 random distractors - making up for a total of 4 pictures per word. The distractors were randomly chosen other objects from the same list. Twelve following trials made up for a block, and the experiment consisted of a total number of 5 blocks. In total, therefore, the whole experiment consisted of 60 (5 x 12) trials. Figure 3 shows a possible setup for the first block of experiment 1.

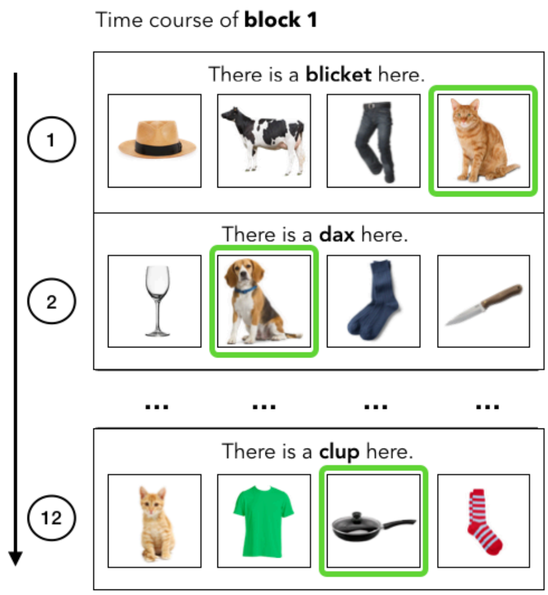

Figure 3: Experimental design of experiment 1 by Chemla & Dautriche

There were two constraints with regard to the objects presented. Firstly, each object would be presented 20 times in total: 5 times as a target - exactly once per block- and 15 times as a distractor in total. Secondly, an object could not be a distractor more than twice for a specific target.

We simulated this experiment by constructing a digital matrix of the 60 trials, in which the distractors per trial were randomized but adhered to the constraints of the original experiment. This matrix was fed to the cross-situational learning and the propose-but-verify models to perform the learning. A crucial difference between the real experiment and our simulated experiment is that Dautriche & Chemla had participants click on the pictures to indicate their choices. Our models do not click on the pictures, which is a point of discussion that we will comment on later (see Discussion).

4.3 Results

Having carried out the simulated experiments, we are interested in the performances of both models. We will evaluate these by means of two types of comparisons. Since we have two models, one for each theory, we can take a closer look at how the obtained results from these models differ from each other. In addition, we can compare the results from each of the models to the data that was obtained in the original Experiment 1 from Chemla & Dautriche.

Model-to-model comparison

We we will now look at results from our two models separately to see if they match our hypotheses and then we will set them side by side to make a comparison between the two.

First, let us look at the results from our **cross-situational learning model**. We chose to run the experiment 5 times during the learning phase, in order boost the models' weights and thus the learning effect. This allows also us to see how the learning trend continues over time. Figure 4 shows us how the connection weights tied to the cues in the model develop over time. We see the process over 300 trials, which is the duration of the entire experiment (60 trials) increased by a factor of 5. We can observe that each time a cue is presented together with the outcome, the weight for that cue increases. We see that the cue competition decreases over time as the weight for the correct cue prevails.

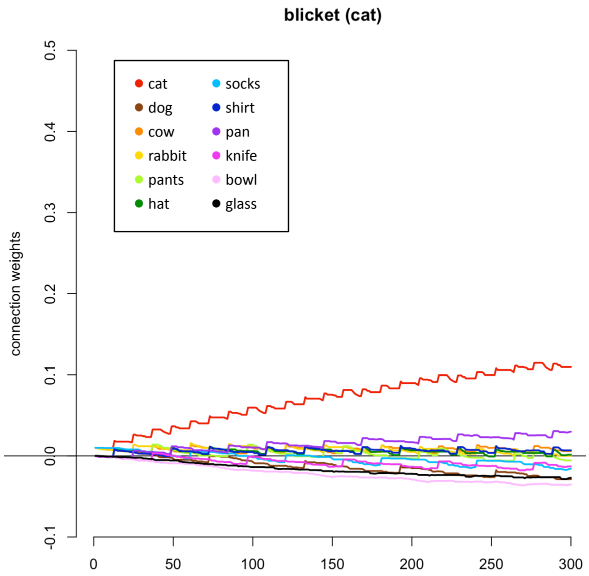

Figure 4: Connection weights for the label "blicket" in experiment 1 using cross-situational learning over 300 simulated trials (25 blocks)

Using this model, we acquire the values of the connection weights that are associated with certain outcomes. We can not simply assume that the cue with the highest association corresponds to the made choice on a trial in the experiment. But since we want to be able to make a comparison between the choices of our two models per trial, we made an interpretation of the weights. In this interpretation, we first summed up the weights corresponding to the four objects that are presented on a trial and consequently divided the weights for each of the cues corresponding to the objects by this total sum. This left us with percentages of activation for each outcome, which we interpret as the choices. Since this is merely our interpretation it does not guarantee that this is how the choices are actually made by the model. This point will be discussed later (see Discussion).

Let us now move on to the performance of the **propose-but-verify model**. This model, as explained before, only keeps track of a single conjecture (its hypothesis) and makes its choice per trial based on that hypothesis with a chance variable, given by the α parameter. The results from this model show us the actual choices that simulated participants made and not just connection weights, as with the cross-situational learning model. We ran the experiment over 1000 simulated participants to gain insight into the average of the choices made by the model. Figure 5 shows us how the accuracy of the choices made per trial develops over the total of 5 blocks.

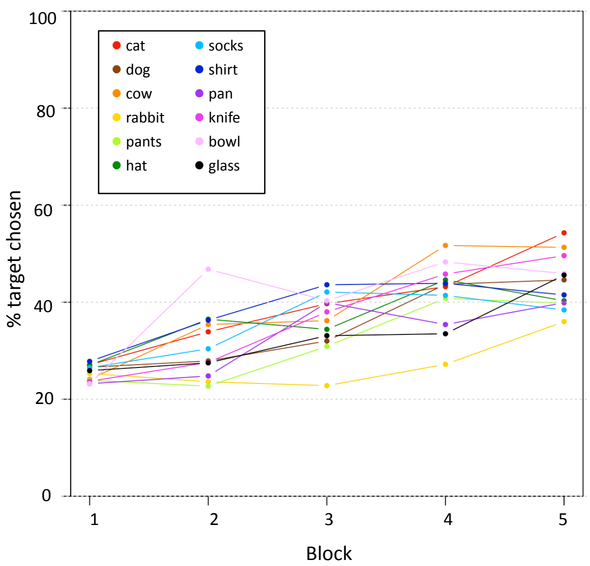

Figure 5: Accuracy per label for each block using our propose-but-verify simulation of experiment 1

We are interested in seeing the parallels and differences between the two just discussed models. For this reason, we decided to focus on one of the labels, "blicket" (cat), and see what choices both of the model make on the trials where this is the target label. To observe the choices made by our cross-situational model, we use the interpretation of the connection weights that we have described earlier. Figure 6 shows us an overview of the five trials (one per block) in which "blicket" is the to-be-learned word. This figure illustrates how often the target and the distractors were chosen by both models. This gives us insight how the models compare to each other. We see the target accuracy in the cross-situational learning model lies higher than that in the propose-but-verify model from blocks 2 to 4. The two reach a similar target accuracy in block 5. Furthermore, we observe that the distribution between the distractors looks similar. It is notable, though, that on the last trial the choice rate for "pan" is bigger in the cross-situational model compared to the propose-but-verify model. What is more, if we look back at figure 4, we see that the connection weight for "pan" in the long run gets a higher activation than all the other distractors. This could hint at the fact that the learning strategy or mechanism changes in the long run.

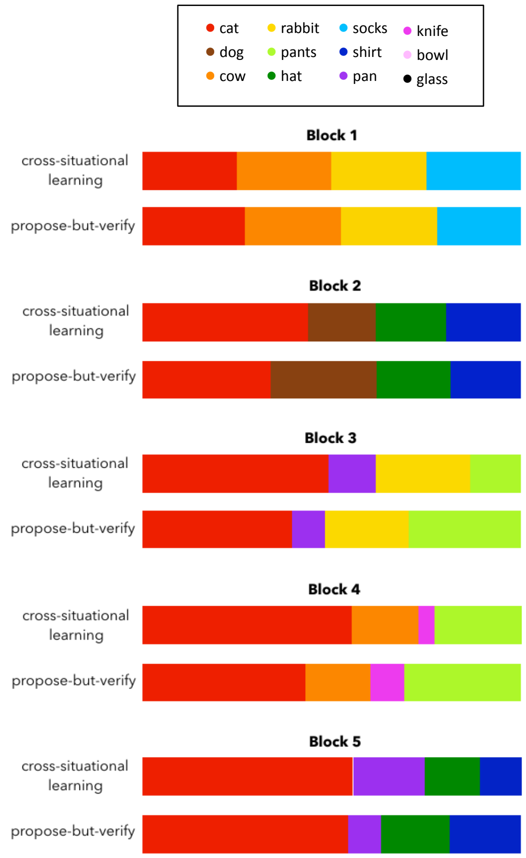

Figure 6: Choices made by the cross-situational learning and propose-but-verify model on trials with "blicket" as target label

Model-to-data comparison

We have taken a closer look at the performances and differences between our two models, but we can also make a comparison between our findings and the data that was obtained in the experiment done with human participants that Dautriche and Chemla carried out. In figure 7 we can observe the difference in the performance of our simulated participants with both our models and the human participants tested by Dautriche & Chemla. All of them start at chance level. Our cross-situational learning model surpasses the accuracy of the human participants quickly in block two and prevails until the last block, where it reaches an accuracy that lies close to the accuracy that is reached by the human participants. But it is noteworthy to mention that accuracy is difficult to extract. Our propose-but-verify model follows a course that instead lies close to the human participants' accuracy until block 4. In the last block, it deviates. It is noteworthy that the learning pattern that can be observed by human participants, seems to lie in-between the patterns that our two models follow. We will discuss a possible explanation for this effect later (see Discussion).

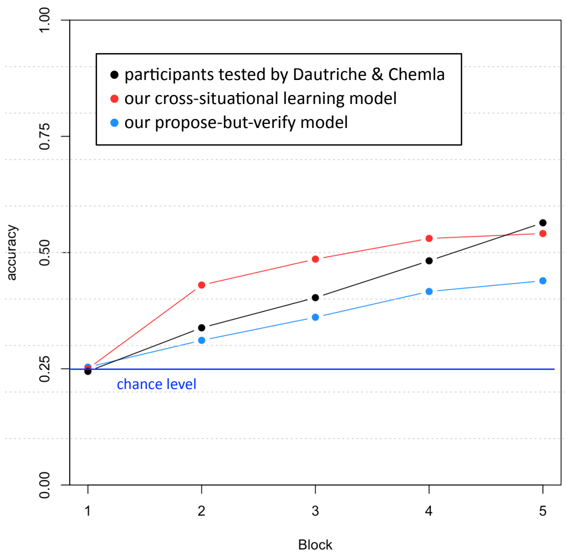

Figure 7: Average accuracy in experiment 1, our models vs. human participants tested by Dautriche and Chemla

5. Experiment 2

5.1 Introduction

Up until now we have only looked at object-word learning in trials with randomized sets of objects. As a result, there was no consistency in contexts between the shown objects. Dautriche and Chemla carried out a second experiment, a modification of their first experiment, such that context starts to play a role. Here, they define context as a consistency in category. An example of such a context could be a zoo context, in which primarily objects belonging to an animal-category are present at the same time. This consistency in context could serve as an aid in object-word learning. We constructed a computational model of Dautriche and Chemla's second experiment, so that we could compare our results of our model of the experiment without the contextual consistency to one where the context plays a role.

5.2 Design

The second experiment that Dautriche & Chemla conducted was overall very similar to the first experiment, in the sense that the same objects and words were used for this experiment. The experimental design did not differ from the first experiment in numbers: the second experiment, again, used 60 trials divided over 5 blocks. But now there is a new (contextual) restriction with regard to which pictures are shown, only in the first block. In this block, the to-be-learned objects were shown together with other objects (distractors) that belonged to the same semantic category. There were three different categories: *animals* (cat, dog, cow, rabbit), *clothes* (pants, hat, socks, shirt) and *dishes* (pan, knife, bowl, glass). Figure 8 shows a possible setup for the first block of experiment 2.

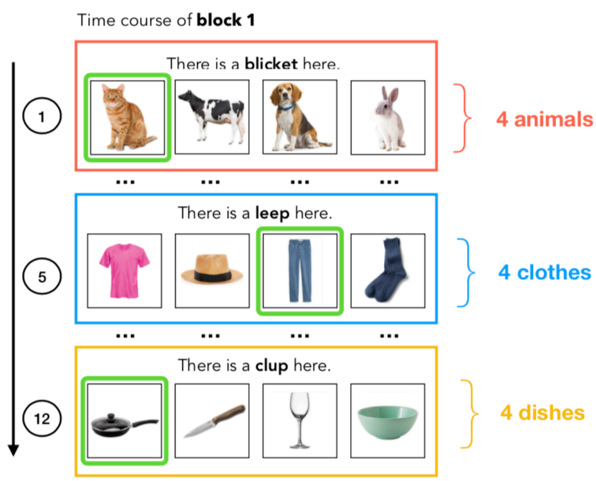

Figure 8: Experimental design of experiment 2 by Chemla & Dautriche

After the first block, the second to fifth block continued without the contextual restriction, i.e. these blocks did not differ in terms of setup compared to experiment 1.

5.3 Results

We have now simulated the second experiment, a modification of the first experiment with added consistent semantic context. We want to inspect to what extent the contextual consistency has served as a learning aid for object-word learning in the performance of our models. We will therefore again make a model-to-model comparison and a model-to-data comparison.

Model-to-model comparison

In an analysis of the results from the first experiment, we looked at how the connection weights of our **cross-situational learning model** developed over time. Figure 9 shows us the connection weights of this model in our simulation of the second experiment. We can not distinguish a clear difference between these weights and those in the first experiment, but we will later interpret these weights as choices made by the simulated participants again, which we can compare to the choices made during the first experiment.

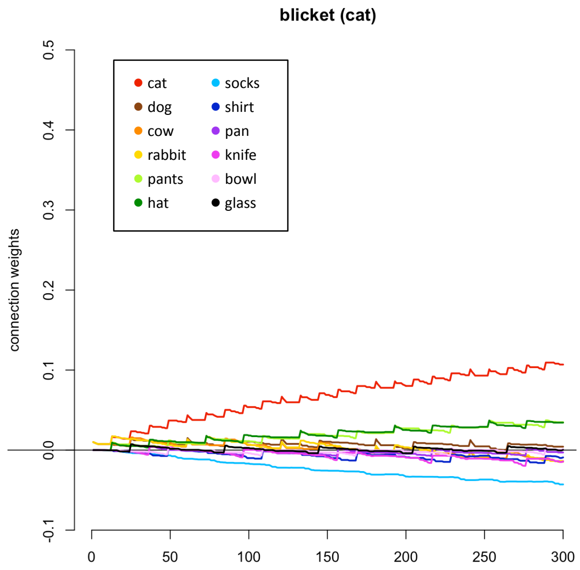

Figure 9: Connection weights for the label "blicket" in experiment 2 using cross-situational learning over 300 simulated trials (25 blocks)

We can also compare the results of our **propose-but-verify model** in the second experiment to the results we obtained in the first experiment. We looked at how the accuracy developed for all labels over the five blocks. We will now look at this development in the second experiment, so that we can see what effect the added consistency in context has had. Figure 10 illustrates the development in accuracy over the five blocks. What stands out from this plot is that for some of the labels the accuracy takes a large jump from the first to the second block, while for other labels the accuracy improves at a much lower rate. We found that the accuracy rate in the second block respective to the first block depends largely on the set of (randomized) distractors that are present in the second block. The participants have seen each of the objects in a semantic coherent context during the first block, so the category of the objects can serve as a memory aid during the second block (e.g. when a participant does not remember which object they chose during the first trial, but if they do remember that it must belong to the category "animal", they will choose an object from that category again). If the distractors belong to different categories than the target, participants are less like to choose them. This is how the distractors influence the accuracy change from one block to the next, in particular from the first to the second block.

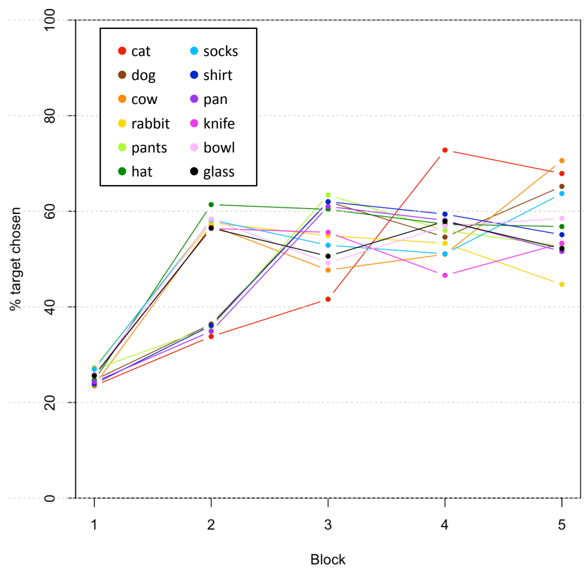

Figure 10: Accuracy per label for each block using our propose-but-verify simulation of experiment 2

Now, we want to compare the models to each other again to gain insight into their similarities and differences. We will again focus on the trials where "blicket" (cat) is the target label and inspect what choices are made by both models on these trials. A visualization of these choices can be observed in Figure 11. Here we can clearly see that distractors that belong to the same category as "blicket", namely animals, are being chosen more often in comparison to other distractors by both of the models. This happens on blocks 2 and 3. Furthermore, we observe that the target accuracy seems very similar between the two models, with only notable difference in the fourth block.

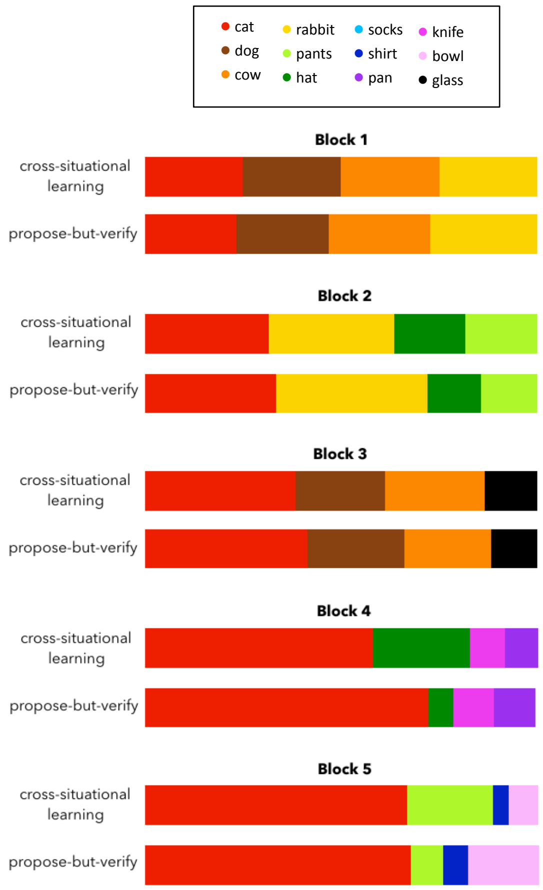

Figure 11: Choices made by the cross-situational learning and propose-but-verify model on trials with "blicket" as target label

Model-to-data comparison

Now that we have seen how the models' performances changed after adding a semantical contextual restriction to the experiment, we can also take a look at how the average accuracy has benefited from this modification in comparison to how the human participants in Dautriche & Chemla's experiment performed. In Figure 12 we can see the difference in the performance of our simulated participants in both of our models and the human participants tested by Dautriche & Chemla. Again, all models start at chance level. The accuracy of our cross-situational learning model stays below the accuracy that human participants obtain until block 3, after which it takes a jump and shows a course that is similar to the data that Dautriche & Chemla obtained. What is remarkable is that if we look at the course of the human data, the patterns is similar to the behavior of our cross-situational learning model during the first experiment. If we look at the results from our cross-situational learning model, we can notice that the pattern lies very close to the data from human participants. Let us then look at the results from our propose-but-verify model. The trend it shows in performance until block 4 lies very close to how the human participants performed. It stands out that the propose-but-verify model and the cross-situational learning model during the first experiment perform similar to human participants until the fourth block, whereas the cross-situational learning model during the second experiments seems to perform similar to human participants in blocks 4 and 5. This might, as we speculated on earlier, hint at the fact that people are in reality using a combination of both strategies.

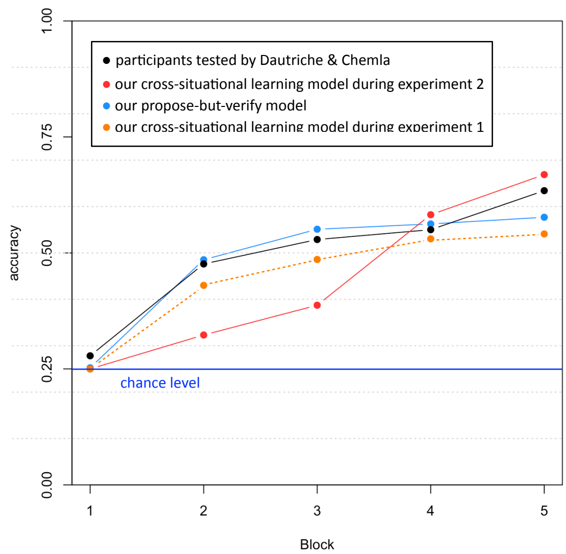

Figure 12: Average accuracy in experiment 2, our models vs. human participants tested by Dautriche and Chemla

It stands out that our cross-situational learning model in blocks 1-3 from the second experiment performs worse than it does during the first experiment, while the addition of contextual consistency serves as an aid for the human participants and the propose-but-verify model. We can see that the performance trend from cross-situational learning model in the first experiment matches the human data better. The difference in these first three blocks can be understood if we take a closer look at the connection weights in trials 1-60 in the first and the second experiment. Figure 13 shows us the connection weights in in experiment 1 the 60 simulated trials, whereas figure 14 shows us the weights in a identical manner but for experiment 2.

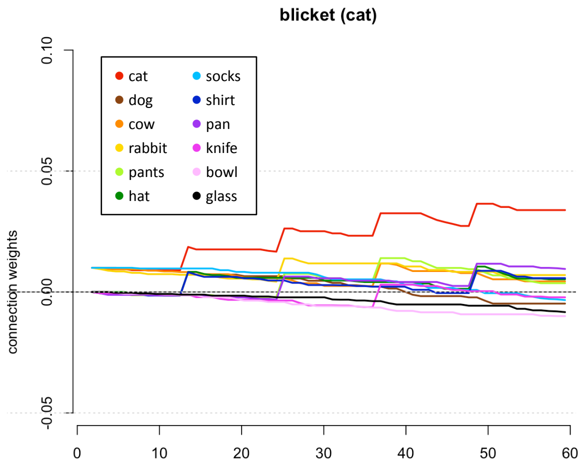

Figure 13: Connection weights for the label "blicket" in experiment 1 using cross-situational learning over 60 simulated trials (5 blocks)

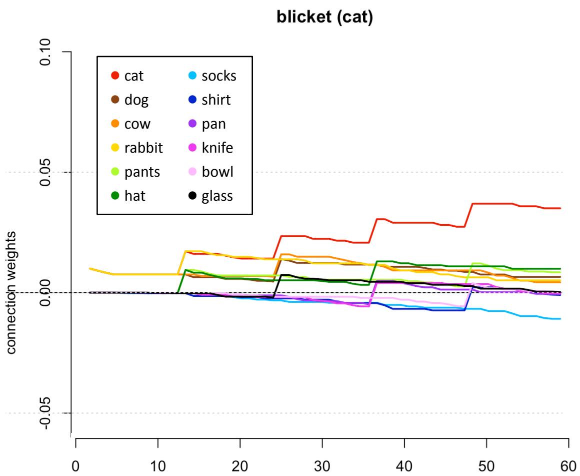

Figure 14: Connection weights for the label "blicket" in experiment 2 using cross-situational learning over 60 simulated trials

When comparing the connection weights in the 60 trials during experiment 1 (figure 13) to those in experiment 2 (figure 14), we can see that due to the addition of context in the second experiment, the weights for cues that belong to the animal-category like "blicket" (cat), compete with the cue for cat until the third block. This is not the case in experiment 1, where the cue for cat starts to prevail after the second block. This explains how the contextual consistency in the first block lowers the performance of the cross-situational learning model in the second experiment as opposed to the first experiment.

Discussion

We constructed computational models of two object-word learning theories, cross-situational word learning and propose-but-verify and used them to test their performances on simulations of experiments done by Dautriche and Chemla in 2014. We can now draw several conclusions from these simulations.

We compared our two models to each other to gain insight in their similarities and differences per block. We found that both show a clear learning trend over time, however we observe minor differences with respect to the competition of distractors for a target.

Furthermore, we made model-to-data comparisons in which we looked at how the performances of our models compared to how human subjects performed in Dautriche and Chemla's experiments. We found that in the first experiment, our cross-situational learning model overall has a slightly higher accuracy than the human participants, while our propose-but-verify model has a slightly lower accuracy. In the second experiment, the accuracy of the propose-but-verify model lies extremely close to the accuracy of human participants until the fourth block. In fact, the accuracy can be approached by taking the combination of the cross-situational model during the first experiment, during the second experiment and the propose-but-verify model.

Our findings from the model-to-model and model-to-data comparisons from both experiments could hint at the fact that in reality, learners might not be using either of the strategies exclusively, but they could in reality be using a combination of both strategies implicitly to obtain a higher level of accuracy in the end. This is what we guess based on looking at the results from the two experiments, though it has been found before by Roembke & McMurray (2016) that a combination of associative learning and real-time processes results in patterns that they have observed in object-word learning during their study.

Some critical points are worth mentioning with respect the choices that we made in order to model the two object-word learning theories

In order to compare the models to each other and to the data obtained by Dautriche & Chemla, we wanted to use the choices made by both of the models per trial so that we could determine the accuracy. However, the cross-situational learning did not return the exact choices made, but it instead provided us with the values for the connection weights for all the cues per trial. In order to be able to make the comparisons, we needed to translate these weights into choices and we did so by interpreting the weights. By using this interpretation, we were able to compare it to the human data, but weights could possibly be interpreted in a different way, which could lead to different insights.

Another point of discussion that we want to make is our approach to model the α-parameter corresponds to the strength of the current hypothesis in the propose-but-verify strategy. In our model, we took a rather crude approach to modeling this parameter. We only used two values, 0.26 and 0.71, that were proposed by Trueswell et al. (2013) in their study on propose-but-verify. We would like to mention that the value of the α-parameter could in reality be of a different value or moreover, more flexible. The value should certainly surpass 0.71 if the experiment is run over a longer period of time, as eventually learners will not forget the meaning of a word.

We would also like to address that in our modeled simulations of the experiments, we only implemented a limited set of cues, which the pictures that were presented on each trial, their category and a background cue. In the experiments carried out with human participants, however, these cues might have been different or more diverse than in the model representation. What is more, the used pictures could have a certain connotation in reality (e.g. someone who really likes cats could be able remember these pictures better), which is too complex to capture in our simulations. This might have had a slight influence on our obtained results.

We found that neither of the two theories that we modeled resembled the human data better than the other. It might be possible therefore that learners in fact use a combination of the mechanisms behind cross-situational learning and propose-but-verify, instead of one of the two exclusively. An earlier conducted study by Roembke & McMurray (2016) also found that a combination of associative learning and real-time processes results in patterns that they observed in object-word learning. However, we have not looked into this combination enough to draw a solid conclusion. It would therefore be interesting to investigate such a combination in a follow-up study. Another element that leaves room for investigation after this study is the modeling of the α-parameter in propose-but-verify. Since we used a rather crude modeling approach of this parameter, we would not have been able to use our propose-but-verify over a longer period of time. This makes it interesting to look into this parameter by trying out different or adaptive rates of forgetting and see how this affects the results and how that compares to how human participants perform.

References

Arppe, A., Hendrix, P., Milin, P., Baayen, R. H., Sering, T. & Shaoul, C. (2015). ndl: Naive Discriminative Learning. R package version 0.2.17. URL: https://CRAN.R-project.org/package=ndl

Bloom, L. (2000). "The intentionality model of word learning: How to learn a word, any word". In Golinkoff, R. M.; Hirsh-Pasek, K.;

Bloom, L.; Smith, L. B.; Woodward, A. L.; Akhtar, N.; Hollich, G. *Becoming a word learner: A debate on lexical acquisition*. New York: Oxford University Press. pp. 19–50.

Dautriche, I. & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(3), 892-903.

Horst, J. S. (2013). Context and repetition in word learning. *Frontiers in Psychology, 4(*149).

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 173.

Kamin, L.J. (1969). Predictability, surprise, attention and conditioning. In B.A. Campbell & R.M. Church (eds.), *Punishment and aversive behavior*, 279–96, New York: Appleton-Century-Crofts.

Miller, Ralph R.; Barnet, Robert C.; Grahame, Nicholas J. (1995). "Assessment of the Rescorla-Wagner Model". *Psychological Bulletin*. American Psychological Association. 117 (3): 363–386.

Quine, WVO. Word and Object. Cambridge, MA: MIT Press; 1960.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Ramscar, M., Dye, M. & Klein, J. (2013). Children Value Informativity Over Logic in Word Learning. *Psychological Science, 24*(6), 1017–1023.

Ramscar, M., Dye, M., Popick, H. M. & O'Donnell-McCarthy, F. (2011). The Enigma of Number: Why Children Find the Meanings of Even Small Number Words Hard to Learn and How We Can Help Them Do Better. *PLoS ONE 6*(7).

Rescorla, R. A, & Wagner, AR. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non reinforcement. In AH. Black & W.F. Prokasy (eds.), Classical conditioning II: current research and theory (pp. 64-99) New York: Appleton-Century-Crofts.

Roembke, T. C. & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language, 87,* 105-127.

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. URL: http://www.rstudio.com/.

Smith, A. D. M. and Smith, K. (2012). Cross-Situational Learning. In Seel, N., editor, Encyclopedia of the Sciences of Learning. pages 864–866. Boston, MA: Springer US.

Smith, K., Smith, A. D. M. & Blythe, R. A. (2010). Cross-Situational Learning: An Experimental Study of Word-Learning Mechanisms. *Cognitive Science, 35,* 480–498.

Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology, 66*, 126–156.

van Rij, J. (2018). NDLvisualization: Additional Visualization Functions for the NDL Framework. R package version 0.2.