The Impact of Lexical Knowledge on Visual Rule Learning in 12- to 14-Month-Old Infants.

Claire van Leeuwen
University of Leiden
Department of MA Linguistics
Theoretical Linguistics and Cognition
Prof. dr. C.C. Levelt
Prof. dr. N.O. Schiller

Table of Contents

1. Introduction

Within the first few months of their lives, children process an astounding amount of information with little to no explicit instruction. The question of how children acquire all the necessary information to generate a clear understanding of their mother tongue – its sounds, syntax, and meaning – has intrigued developmental psychologists and linguists for years.

Researchers have questioned whether children use syntactic knowledge – understanding the rules that govern the formation of sentences in a language – when constructing their first utterances, or if they depend on simple ordering rules (Crain, 1999). If children depended only on simple ordering rules, then this would imply that children possess a completely different system for language than adults, and that at a certain point in their development, children switch over from this simple system to the more syntactic system that we know adults use. However, the hypothesis is that children possess the same competence as adults (Borensztajn, 2009). This theory is strengthened by research by Valian (1986), which looks into the placement of determiners in the speech of English speaking children. This research showed that children between the ages of two and two and a half made no errors in the placement of determiners. Though some children omitted the copula 'be' in their utterances, the self-correction of another child indicates that the children realized that this form is incorrect in English. Self-correction also occurred in substitution tests regarding adjectives, nouns and prepositions. Thus, the research by Valian showed that children make use of syntactic categories and phrase structure rules. This further strengthens the theory that from their earliest utterances, children's grammars are syntactically based (Crain, 1999). Crain argues that these results indicate that infants and adults share the same system for acquiring language via a syntactic system, meaning that they form an understanding of the rules that govern word ordering in sentences in a particular language. When children do construct sentences that deviate from adult-like sentences, these competence errors are based on a lack of sufficient positive evidence for a particular structure (Crain, 1999).

These results made researchers question what input aids infants in creating a fully functional language system. Besides the acquisition of the sound system (Bijeljac-Babic et al., 1993; Maye et al., 2002; Aslin, Jusczyk, & Pisoni, 1998) and, eventually, the lexicon (Benavides-Varela & Mehler, 2014), infants acquire a grammatical system. However, researchers questioned what type of learning mechanism infants use in acquiring their grammatical system.

Saffran et al. (1996) argued that infants possess acute statistical abilities that aids them in grammar acquisition. They showed that infants use their acute statistical abilities to successfully segment phrases using the transitional probabilities of the language. When presented with a continuous stream, the infants had longer looking times towards nonwords, words that violated the statistical probabilities. The infants could extract the necessary information about sequential statistical probabilities after only two minutes of exposure. Further evidence that infants also rely on the statistical distributional properties in syntax comes from Gomez and Gerken (1999). While Saffran et al. (1996) demonstrated that infants can segment words in phrases when the statistical transitional probability within a word in 1.0, Gomez and Gerken argue that transitional probabilities between words in natural sentences almost never equals 1.0. Therefore, they tested whether 11- and 12-month-old infants were able to distriminate between grammatical and ungrammatical strings after only two minutes of exposure to the grammar. They tested the syntactical knowledge of the infants

Marcus et al. (1999) argues that, though these experiments show a learning of sequences, the habituation stimuli and the testing stimuli were the same (e.g. Saffran *et al.,* 1996; Gomez and Gerken, 1999). Therefore, the infants might have depended on transitional probabilities rather than having acquired a rule about the grammar. For example, while an infant might look longer towards a three-syllable word when he or she was habituated with a two-syllable word this might not indicate that the infant has realized there is a violation of the rule. Rather, the infant might look longer because he or she has applied a statistical system to rule that the three-syllable word has, on average, more syllables than the preceding utterances. Marcus (1999) also argues that, though the infants in Gomez and Gerken (1999) could distinguish between grammatical and ungrammatical sentences, this does not necessarily imply rule learning. For example, the infants could have used statistical information about the transitional probabilities, for example X is never followed by Y in this grammar. The research by Marcus et al. (1999) questioned whether infants showed similar results – that the infants attended longer towards novel grammatical instances – when the testing and habituation stimuli were different. In their experiment seven-month-old infants were habituated to artificial three-word sentences in either an ABA or an ABB grammar, and then tested with novel three-word sentences, which either consisted of the same grammar or a inconsistent grammar. For example, an infant habituated with the ABA grammar was exposed to pharses such as “le di le,” “le je le,” “le li le,” “le we le,” and then tested with “ba po ba,” “ko ga ko,” (consistent with ABA) “ba po po,” and “ko ga ga” (consistent with ABB). Because the test items did not appear in the habituation phase, the infants could not use statistical probabilities to differentiate between grammatical and ungrammatical. Instead, they would have to rely on rule learning. The infants had a longer looking time during the presentation of the inconsistent sentences, which demonstrated that they could differentiate between the consistent and the inconsistent grammar through rule learning. These results indicate that infants can generate abstract grammatical rules and apply these to novel instances.

An issue that arises from these experiments is whether the mechanisms that govern these statistical and rule learning abilities are domain specific (i.e. language specific), or if they operate more generally. Various experiments have already demonstrated that infants can track sequential statistical information in both linguistic (Saffran et al., 1996) and non-linguistic material such as musical tones and visual shapes (Saffran et al., 1999; Kirkham et al., 2002). However, infants behave differently in rule learning experiments. Though infants are able to generate a rule when presented with linguistic stimuli (Marcus et al., 1999), they fail to generate a rule about non-linguistic stimuli, even if the stimuli is similar to those used in the statistical learning experiments. Therefore, Marcus et al. (2004) argues that this rule learning mechanism is language specific.

However, Saffran et al. (2007) argues that previous research had not reconciled the different important features in non-linguistic and linguistic stimuli. These features, such as familiarity and categorizability, can influence an infants’s ability to generalize rules. Since infants are already familiar with linguistic stimuli, and because the stimuli is all easily identifiable as part of the same set (i.e., linguistic items), they are easily categorizable. However, when infants are presented with non-linguistic items, such as geometric shapes, they might not see them as members of the same category and are, therefore, unable to generate a rule. To test this hypothesis, Saffran et al. replicated the Marcus et al. (1999) experiment with visual stimuli instead of linguistic stimuli. They used pictures of dogs for the visual stimuli, as many infants are already familiar with dogs and, although they might look slightly different from one another, infants can categorize them. Saffran et al. put forth the hypothesis that the difficulty infants have in discriminating other input, such as musical notes or geometric shapes, might be caused by their inability to categorize these stimuli as belonging to the same group. According to Saffran et al. this causes every triad to appear completely unique to the infants, rather than being an exemplar of the general pattern. They conclude that infants are able to generalize rules when the stimuli are suited to specific rule learning mechanism; it does not matter whether these stimuli are linguistic or non-linguistic.

In this experiment, we hypothesize that an infant's ability to generalize rules about non-linguistic (i.e., visual) stimuli is impacted by their lexical knowledge of the presented item. Thus, the stimuli do not have to be in the same category (e.g., dogs). It might already be sufficient if the infant possesses a lexical understanding of the object. Thus, if infants have a lexical understanding of the object, then it should be easier for them to categorize the presented stimuli. However, if they lack this understanding, or if they are presented with nonsense visuals, they should be unable to categorize these images and it will negatively impact their ability to generalize a grammatical rule. To test this hypothesis, we replicated Saffran et al.'s (2007) experiment, except the stimuli used were not pictures of dogs. Rather, we used visuals of inanimate objects that we assumed the infants to be familiar with. To see whether lexical knowledge was sufficient information for the infants to generalize a rule, we also tested an unknown condition, with items we assumed the infants to be unfamiliar with, and a nonsense condition, with abstract visuals the infants could not have lexical knowledge of. We chose to use inanimate objects to ensure that the infants would not look longer towards the animate object, simply because they find them more interesting.

2. Method

2.1. Participants.

Participants were 59 full term infants aged between twelve and fourteen months of age (M age = 13;37, SD = 0,83). Sixty infants were tested, but due to premature birth, one of the male infants was excluded after the experiment was concluded. Because of time constraints, no other male infant was tested. All infants were counterbalanced for gender and condition. The infants were assigned randomly to one of the three conditions and one of the two grammars, and gender was balanced.

None of the infants had hearing difficulties, visual impairments or were in any other way developmentally delayed. Other infants were excluded from the final sample due to fussiness (9 infants), experimentor errors (7 infants), premature birth (5 infants), both parents being dyslexic (1 infant), falling asleep (1 infant), and parental interference (1 infant).
The participants were recruited from the existing Baby Lab database via letter and telephone. Parents and infants received a small gift (a picture book) for their participation.

	Known	Unknown	Nonsense
Bilingual	4	8	8
Monolinual	15	12	12

Table 1. Distribution of bilingual and monolingual infants per condition

2.2. Apparatus and Stimuli.

The infants were tested in one of the three conditions: known, unknown and nonsense. The items in the known condtions were items that the infants should be highly familiar with and that should be common. The items in the unknown condtions should be items that the infants might have seen, but with which they were not familiar and that were not common. In order to find which words the infants would know, or would be familiar with, the *Lexilijst Nederlands* was consulted. This lists the percentage in which children between 15 and 27 months of age know a word. This list is standardized on a sample of 809 infants, representative of the Netherlands. Using this list, 12 words per condition were gathered. The criteria for which words were used in the two conditions were the reported percentage of children who knew a particular word, and the words had to represent an inanimate object; for the known condition the percentage was at least 55%, and for the unknown condition the percentage was below 29%. Table 1 shows the items and their corresponding percentage.

Known words (percentage)		Unknown words (percentage)	
Ball (83)	Book (61)	Umbrella (29)	Balloon (25)
Car (76)	Shoes (60)	Carrot (17)	Pencil (26)
Bath (69)	Sock (60)	Leaf (19)	Candle (27)
Coat (68)	Bicycle (58)	Cup (20)	Crayon (12)
Apple (64)	Banana (55)	Helicopter (23)	Lego (18)
Bed (63)	Chair (55)	Plant (24)	Chocolate (27)

Table 2. Known and Unknown words

Additional checking of the infants' lexical knowledge was done using the Dutch Communicative Development Inventories (N-CDI 1) list (Zink & Lejaegere, 2002), with an addendum containing the words used in the unknown condition but which were not in the original N-CDI questionnaire. The results of this questionnaire can be found in table 3 and table 4. Table 3 shows the items used in the known condition and the reported number of infants per condition that knew these items. Table 4 shows the items used in the unknown condition and the reported number of infants per condition that knew these items. The items used in the test trials are in bold.

Known items	known	unknown	nonsense	all	%
ball	6	10	1	17	28
car	9	10	7	26	44
bath	11	11	6	28	47
coat	16	12	7	35	59
apple	8	13	2	23	38
bed	6	11	7	34	57
book	10	6	2	18	30
shoes	10	12	3	25	42
sock	**12**	**11**	**6**	**29**	**49**
bicycle	**7**	**9**	**1**	**17**	**28**
banana	**9**	**10**	**5**	**24**	**40**
chair	**7**	**9**	**0**	**16**	**27**

Table 3. Results N-CDI questionnaire for comprehension known stimuli

unknown items	known	unknown	nonsense	All	%
balloon	9	9	4	22	37
leaf	1	4	**0**	**5**	8
candle	1	3	0	4	6
cup	0	6	0	6	10
crayon	1	1	0	2	3
umbrella	0	1	0	1	1
plant	3	5	0	8	13
pencil	1	5	1	7	11
chocolate	**1**	**3**	**0**	**4**	**6**
helicopter	**0**	**5**	**1**	**6**	**10**
lego	**1**	**3**	**1**	**5**	**8**
carrot	**2**	**4**	**0**	**6**	**10**

Table 4. Results N-CDI questionnaire for comprehension unknown stimuli

Table 3 and table 4 show that the infants' reported comprehension of the stimuli in both the known and the unknown condition can deviate for the numbers presented in the *Lexilijst Nederlands* (table 2). In table 3, three of the items are only known by 37% of the infants, two of which are part of the testing stimuli.
Table 4 also shows a deviation with the 'balloon' item; this word was reported to be known by 44% of the infants, a difference of 20% between *Lexilijst Nederlands* and the reported number in the N-CDI questionnaire.
For each of these words, corresponding high resolution images were found and displayed on a white background. Twelve additional pictures were used for the nonsense condition, these pictures were abstract and unrecognizable (figure 1).

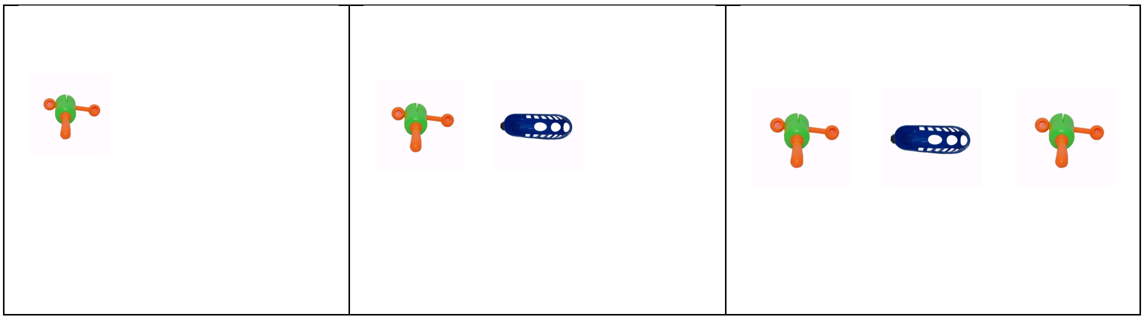

Figure 1. triad ABA Nonsense "greenorange-blue-greenorange"

The images were taken from Google images and the NOUN database (Horst & Hout, in press). Using PowerPoint, a triad was created using the high-definition pictures on a white background in order to ensure that no other visual information was present (figure 2).

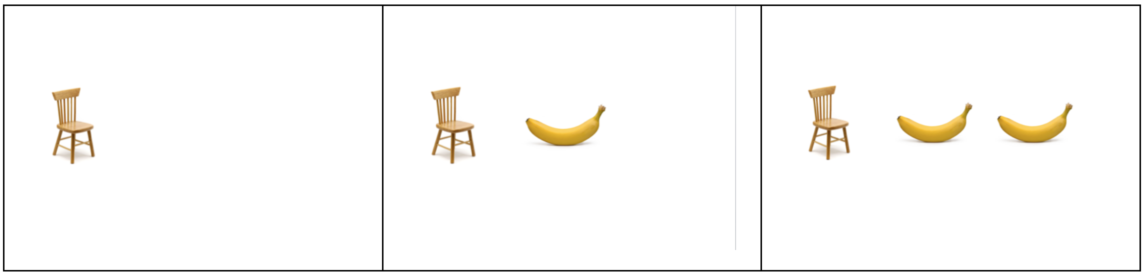

Figure 2. Triad "chair-banana-banana"

The triads were created by combining items from the A-category and the B-category, which can be found in table 5. Therefore, a triad in the ABA grammar consisted of an item from the A-category, followed by an item from the B-category, which is followed by a repition of the initial (A-category) item. For example, a triad in the Unknown condition for the ABA grammar during the familiarization phase can be "plant-balloon-plant".

Known		Unknown		Nonsense	
A words	B words	A words	B words	A words	B words
Apple	Car	Plant	Balloon	Greenorange	Blue
Ball	Bath	Leaf	Candle	Grey	Pink
Bed	Book	Cup	Crayon	Purpleturquois	Q-tip
Shoes	Coat	Umbrella	Pencil	Wire	Slapper

Table 5. Stimuli used in familiarization per condition

The triads used during the testing phase were created by combining four novel items (table 6). Thus, a triad in the testing phase of the Unknown condition can be "carrot-lego-carrot" (when consistent with the ABA grammar) or "carrot-lego-lego" (when consistent with the ABB grammar).

Known		Unknown		Nonsense	
A words	B words	A words	B words	A words	B words
Chair	Banana	Carrot	Lego	Pokemon	Virus
Bicycle	Sock	Helicopter	Chocolate	Bill	Galaxy

Table 6. Stimuli used in the testing phase per condition

2.3. Procedure

Infants were tested individually and sat on their caregiver's lap in a sound-attenuated room with a low light intensity. The caregiver wore glasses covered with duct-tape to make them blind to the stimuli and to reduce parental interference. The infants were seated one meter away from the screen. An attention grabber would play on the screen until the infant centered their attention. The attention grabber consisted of a loop of bubbles on a purple background with the sound of a harp playing. Once the infant centered his or her attention, the familiarization phase begun. During this phase, which consisted of four phases of 30 seconds, the infants were familiarized with one of the two grammars (ABA or ABB) in one of the three conditions (known, unknown or nonsense). The triads were randomized for each of the condition, in both the ABB and the ABA grammar. For example, a triad in the known condition could be "chair-banana-banana". The first picture (e.g., a chair) appeared on the left of the screen for 0.3 seconds. After 0.3 seconds, the second picture (e.g., a banana) appeared in the middle of the screen together with the first picture for 0.3 seconds. After 0.3 seconds, the last picture appeared on the right on the screen for 0.8 seconds. After this, the triad started anew, with a maximum of 7 repititions with a maximum duration 15 seconds (see figure 2).

Between each repition of the familiarization phase and between each trial, the attention grabber was played to attract the attention of the infants. When the infant focused his or her attention on the screen for more than two seconds, the attention grabber would stop playing and the familiarization or trial would start. If the infant looked back within the two-second margin, then the trial continued.

The stimuli were randomized for each infant. An examiner outside of the sound-attenuated booth followed the infant's behavior on a monitor. The looking behavior was then coded offline, using ELAN[1] (Sloetjes & Wittenburg, 2008) and analyzed in SPSS[2].

2.4. Coding

The looking behavior of the infants was recorded and then analyzed offline, using ELAN. In ELAN the trials were coded on four tiers, which consisted respectively of the looking during trial, looking away during trial, looking away during attention grabber, and looking during attention grabber. If an infant looked away and looked back within the two-second cut-off, looking times were coded as, for example, for trial 1; 1a, 1b, 1c, etc. This would then indicate that the infant has looked away and back again three times within one trial. These looking times was added up in Excel, so that if an infant looked away three times during a trial, the sum of the looking time was used in the analysis. An infant was considered to look when both of their eyes were focused on the screen (i.e. in the direction of the camera).

The data was exported to and gathered in Excel for each infant, per trial, per condition, and overall.

Infants were excluded if they were fussy, crying or sleeping, or if their caregiver intervened during the testing. A note of this was made, and the results were excluded after the experiment was finished.

3. Results

For the first analysis, we used the data of all 59 successful infants (*M* age = 13;37 months of age, SD = 0,83). One infant was excluded after testing was finished, because of his premature birth. Due to time constraints no new infant was tested. The N-CDI results can be seen in table 6 (M = 24; SD = 6,499) and table 7 (M = 6,3; SD = 5,31).

[1] Retrieved from http://tla.mpi.nl/tools/tla-tools/elan/. At Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands

[2] IBM Corp. Released 2012. IBM SPSS Statistics for Macintosh, Version 21.0. Armonk, NY: IBM Corp.

Known items	All	%	Unknown items	All	%
ball	17	29	Balloon	22	37
car	26	44	Leaf	5	8,5
bath	28	47	Candle	4	6,8
coat	35	59	Cup	6	10
apple	23	39	Crayon	2	3,4
bed	34	58	Umbrella	1	1,7
book	18	31	Plant	8	14
shoes	25	42	Pencil	7	12
sock	**29**	**49**	**Chocolate**	**4**	**6,8**
bicycle	**17**	**29**	**Helicopter**	**6**	**10**
banana	**24**	**41**	**Lego**	**5**	**8,5**
chair	**16**	**27**	**Carrot**	**6**	**10**

Table 7. N-CDI results for tested infants for known and unknown items

Table 7 shows that the infants are reported to know less words than initially inticipated using the *Lexilijst Nederlands*. However, the items that are part of the known stimuli are reported to be understood by a larger number of the infants than the items in the unknown condition.

Figure 3 shows that there are no outliers in the overall distribution of Looking Times. Therefore, no further infants needed to be exluded due to lack of looking time.

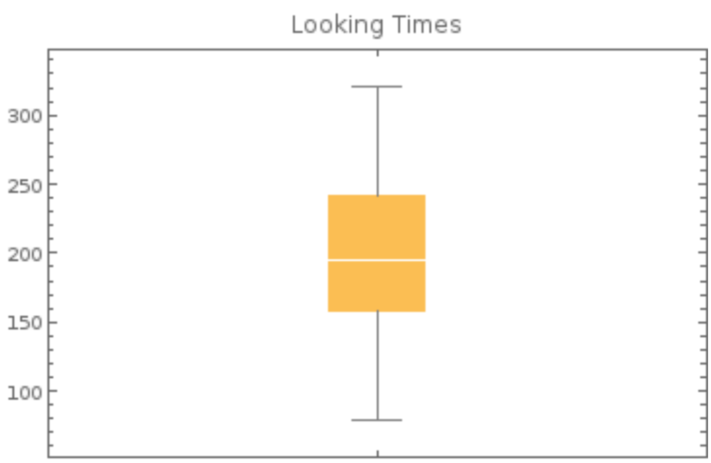

Figure 3. Distribution of overall Looking Times.

The data was analyzed using a univariate linear model in SPSS, with the dependent variable being the mean looking time per block for both the consistent and the inconsistent grammar. The other fixed factors were: gender, bilingualism, condition (known, unknown or nonsense), training grammar (ABA or ABB), and consistency (consistent or inconsistent).

First, a between subject ANOVA was conducted to compare the effects of gender and bilingualism on looking times. Neither gender and nor bilingualism had a significant effect on looking times [F(1) = 0.166, p = 0.684] (gender) and [F(1) = 0.473, p = 0.492] (bilingualism). Since these factors did not have a significant effect on the looking times, they were taken out in the following test.

Another between subject ANOVA was conducted to compare the effects of condition, training grammar and consistency on the looking times. There was a significant main effect of condition between the three different conditions: known, unknown and nonsense at the $p<.05$ level for these three conditions [F(2) = 5.527, p = 0.004]. The mean score for the known condition (M = 10.034, SD = 0.453) is significantly different to the unknown condition (M = 8.334, SD = 0.441) and the nonsense condition (M = 8.099, SD = 0.441). These results suggest that the infants attended significantly longer in the known condition than in the unknown and nonsense condition. There is also a significant interaction between condition and training grammar [F(2) = 4.676, p = 0.01] with the ABA training grammar (M = 8,566, SD = 0,367) and the ABB grammar (M = 9,078, SD = 0,360) (Figure 4). However, there was no main effect of the training grammar on the looking times [F(1) = 0.994, p = 0.320).

Mean listening times (s) (SD)			
condition			Repeated measures analysis of variance
	consistent	**inconsistent**	
Known	9,989 (0,641)	10.079 (0,641)	F (1) = 2,333, P = 0,145
Unknown	8.24 (0,624)	8.43 (0,624)	F (1) = 0,440, P = 0,516
Nonsense	8.596 (0,624)	7.601 (0,624)	F (1) = 1,106, P = 0,308

Table 8. Mean looking time to consistent and inconsistent for each condition.

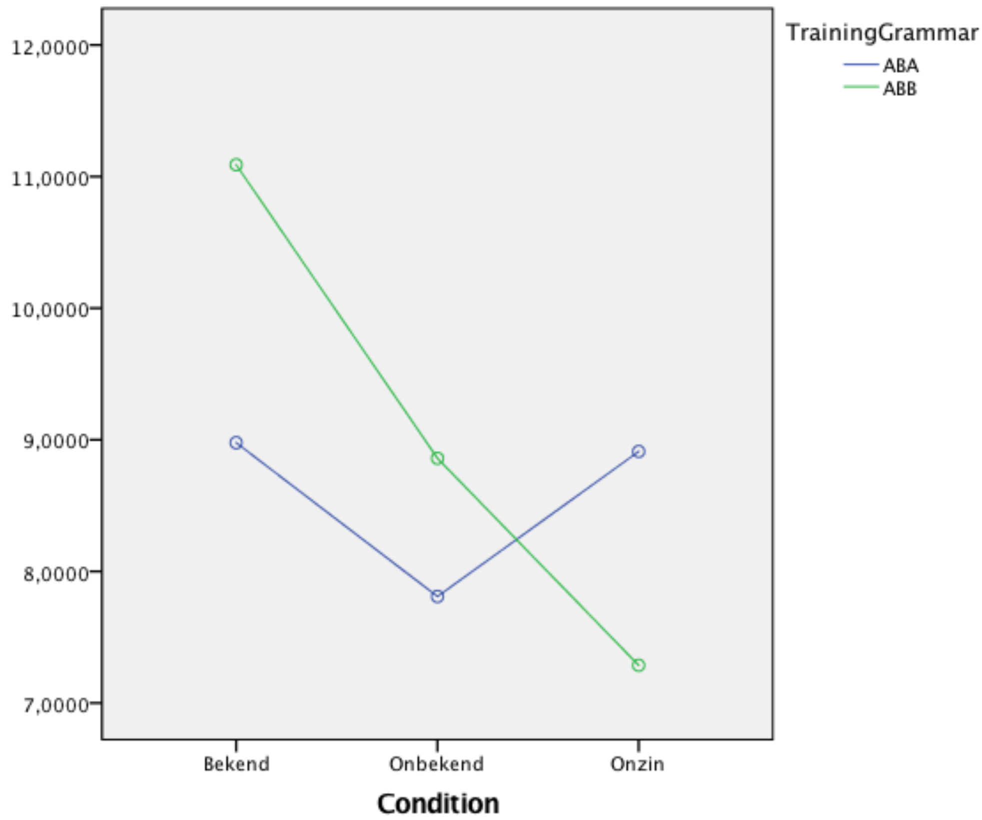

Figure 4. Interaction between Condition nd Training Grammar

The above results called for a separate analysis of the three different conditions.

3.1 Known

Nineteen full-term infants (10 females, 9 males) with no reported hearing difficulties or other developmental delays, composed the final sample (*M* age = 13;42, SD = 0,88). Fourteen additional infants were tested but excluded from the final sample due to premature births (6 infants), fussiness (4 infants), experimentor errors (3 infants) and both parents being dyslexic (1 infant). Table 9 shows the number of infants who were reported to know the stimuli in the N-CDI questionnaire (M = 10,083, SD = 3,12).

Known items	known	%
ball	16	31
car	9	47
bath	11	57
coat	10	84
apple	6	42
bed	8	84
book	16	52
shoes	10	52
sock	**9**	**63**
bicycle	**7**	**36**
banana	**12**	**47**
chair	**7**	**36**

Table 9. N-CDI scores in the Known condition

The data in table 9 shows that the reported numbers of infants that knew the items used as stimuli in the known condition is lower than the numbers reported in the *Lexilijst Nederlands*. This may have negatively effected the infants' ability to differentiate between the stimuli and to generate a grammatical rule.

The data was analyzed using a Repeated Measures ANOVA to compare the effects of block, training grammar and consistency on the looking times. Within-subject factors were block and consistency, the between-subject factors was the training grammar. The mean looking times were per block and per condition (3x2). Training grammar was found to not have a significant main effect on looking times [F(1) = 2,333, p = 0,145, ABA training grammar (M = 8,978, SD = 1,003), ABB training grammar (M = 11,090, SD = 0,952)]. Consistency also had no signifant main effect on looking times [F(1) = 0,015, p = 0,903), consistent (M = 9,989, SD = 0,910), inconsistent (M = 10,079, SD = 0,627)].

Only block was found to have a significant effect on the looking times [F(2) = 4,593, p = 0,017, block 1 (M =11,704, SD = 0,683), block 2 (M =9,354, SD = 0,989), block 3 (M = 9,042, SD = 0,955)]. These results indicate that per block of trials the looking time decreased significantly (see figure 5).

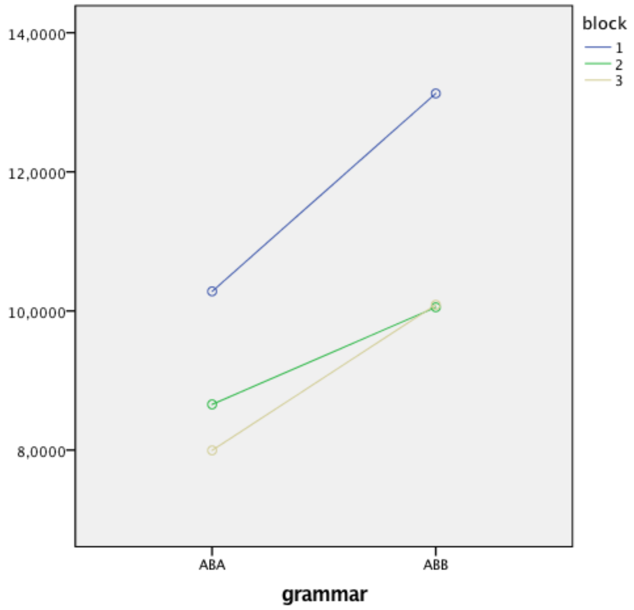

Figure 5. Looking time per block in the known condition per training grammar

Figure 5 also shows that looking times for the ABB grammars were higher than the looking times for the ABA grammar. However, this difference is not significant [F(1) = 2,333, p = 0,145, ABA (M = 8,978, SD = 1,003), ABB (M = 11,090, SD = 0,952)]. Therefore, these results do not indicate learning of the familiarized grammars. However, there is a trend for longer looking times for the infants trained with ABB, as well as a preference for ABB for both the infants trained in ABA and in ABB.

There were no significant interactions. The interactions between block and training grammar [F(1) = 0,141, p = 0,712], between consistency and training grammar [F(1) = 2,209, p = 0,155], between consistency and block [F(2) = 0,203, p = 0,817], and between consistency, block, and training grammar [F(2) = 2,902, p = 0,069] did not yield any significant results.

Bilingualism was analysed using a Repeated Measures ANOVA. The distribution of bilingualism between the infants in the known condition can be seen in table 8. The factor bilingualism did not have a significant effect on looking times [F(1) = 0,182, p = 0,675), bilingual (M = 10,694, SD = 1,596), monolingual (M = 9,928, SD = 0,824)] (see table 1).

Gender was also analyzed using a Repeated Measures ANOVA to see if this factor had a significant effect on looking times. The distribution can be seen in table 9. This factor also did not have a significante effect on looking times [F(1) = 2,993, p

= 0,102] with Female (M = 11,203, SD = 0,936) and Male (M = 8,851, SD = 0,986)

3.2 Unknown

Twenty full-term infants (10 females, 10 males) with no reported hearing difficulties or other developmental delays, composed the final sample (M age = 13;2, SD = 0,72). Five additional infants were tested but excluded from the final sample, due to fussiness (2 infants) and experimenter errors (3 infants).

Table 10 shows the number or infants reported to know the items used as stimuli in the unknown condition in the N-CDI questionnaire (M = 4,08, SD = 2,19).

Unknown items	known	%
Balloon	9	45
Leaf	4	20
Candle	3	15
Cup	6	30
Crayon	1	5
Umbrella	1	5
Plant	5	25
Pencil	5	25
Chocolate	**3**	**15**
Helicopter	**5**	**25**
Lego	**3**	**15**
Carrot	**4**	**20**

Table 10. Reported N-CDI scores infants unknown items

The data was analyzed using a Repeated Measures ANOVA to compare the effects of block, training grammar and test grammar on the looking times. Within-subject factors were block and consistency, the between-subject factor was the training grammar. Looking times were the mean looking times per block and per consisntency (3x2). Training grammar did not have a significant effect on the looking times [F(1) = 0,440, p = 0,516], ABA grammar (M = 7,809, SD = 1,120), ABB grammar (M = 8,859, SD = 1,120). Consistency was shown to be nonsignificant [F(1)

= 0,097, p = 760), consistent (M = 8,240, SD = 0,884), inconsistent (M = 8,428, SD = 0,810)].

As in the known condition, only block was found to be significant [F(2) = 7,419, p = 0,002, block 1 (M = 10,256, SD = 0,903), block 2 (M = 8,016, SD = 0,891), block 3 (M = 6,731, SD = 1,063)]. These results indicate that per block of trials the looking time decreased significantly (see figure 6).

The interactions between block and training grammar [F(2) = 1,581, p = 0, 220], between consistency and training grammar [F(1) = 0,151, p = 0,702], between consistency and block [F(2) =1,074, p = 0,352], and between consistency, block, and training grammar [F(2) = 1,006, p = 0,376] did not yield any significant results.

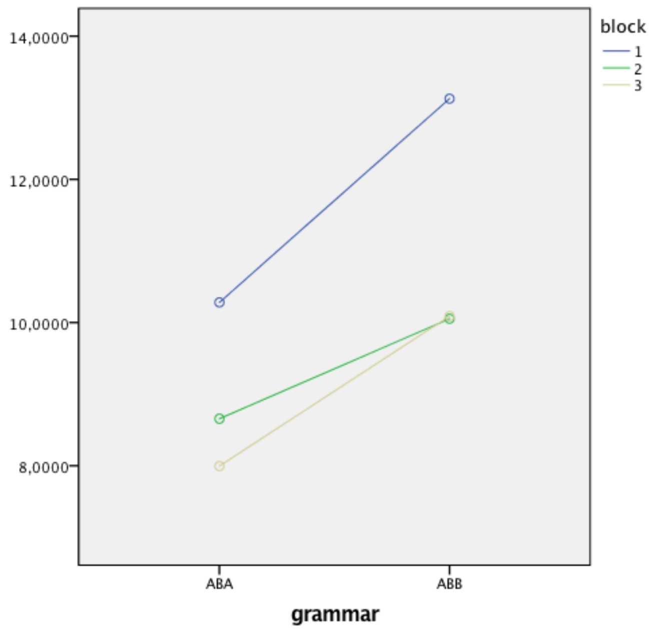

Figure 6. Looking time per block in the unknown condition per training grammar

Similar to figure 5, figure 6 also shows that looking time is higher for the infants familiarized with the ABB grammar than those familiarized with the ABA grammar. However, this difference is not significant [F(1) = 0,440, p = 0,516, ABA (M = 7,809, SD = 1,120), ABB (M = 8,859, SD = 1,120)]. Therefore, these results do not indicate learning of the familiarized grammars.

Bilingualism was analysed using a Repeated Measures ANOVA. The distribution of bilingual infants in the unknown condition can be seen in table 1.

Bilingualism did not have a significant effect on the looking times [F(1) = 0,479, p = 0,498, bilingual infants (M = 9,005, SD = 1,251), monolingual infants (M = 7,887, SD = 1,021)].

Gender was also analyzed using a Repeated Measures ANOVA to see if this factor had a significant effect on looking times. The distribution can be seen in table 12. This factor did not have a significante effect on looking times [F(1) = 1,969, p = 0,178], female (M = 7,266, SD = 1,076), male (M = 9,402, SD = 1,076)].

3.3. Nonsense

Twenty full-term infants (9 females, 11 males) with no reported hearing difficulties or other developmental delays, composed the final sample (*M* age = 13;47, SD = 0,85). Seven additional infants were tested but excluded from the final sample, due to fussiness (five infants), parental interference (1 infant) and experimentor errors (1 infant).

The data was analyzed using a Repeated Measures ANOVA to compare the effects of block, training grammar and test grammar on the looking times. Both block and consistency had a main significant effect on looking time. Block had a significant effect [F(2) = 3,356, p = 0,046, block 1 (M = 9,182, SD = 1,125), block 2 (M = 8, 074, SD = 1,006), block 3 (M = 7,040, SD = 1,010)]. Figure 7 shows the decrease of looking time per block. Compared to the looking times per training grammar shown in figure 5 and figure 6, the looking times per grammar have reversed in the nonsense condition. Figure 7 shows a longer looking time towards to ABA grammar than towards the ABB grammar. However, this difference is not significant [F(1) = 0,757, p = 0,396, ABA (M = 8,911, SD = 1,320), ABB (M = 7,286, SD = 1,320)].

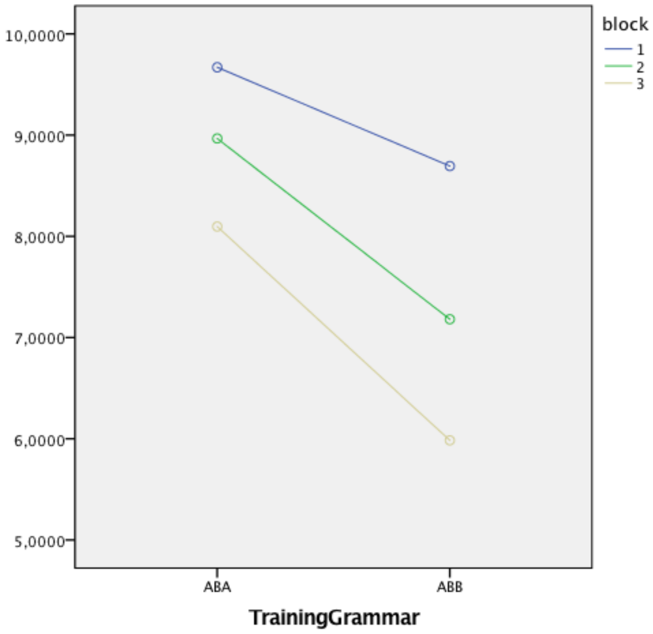

Figure 7. Looking time per block in the nonsense condition.

However, consistency did have a significant effect on looking time [F(1) = 5,327, p = 0,033, consistent (M = 8,596, SD = 0,968), inconsistent (M = 7,601, SD = 0,948)]. This indicates that there was learning. Figure 8 shows that with both the ABA and the ABB training grammar, looking time is significantly longer towards the consistent grammar. This is what we had hypothesized to happen in the known condition.

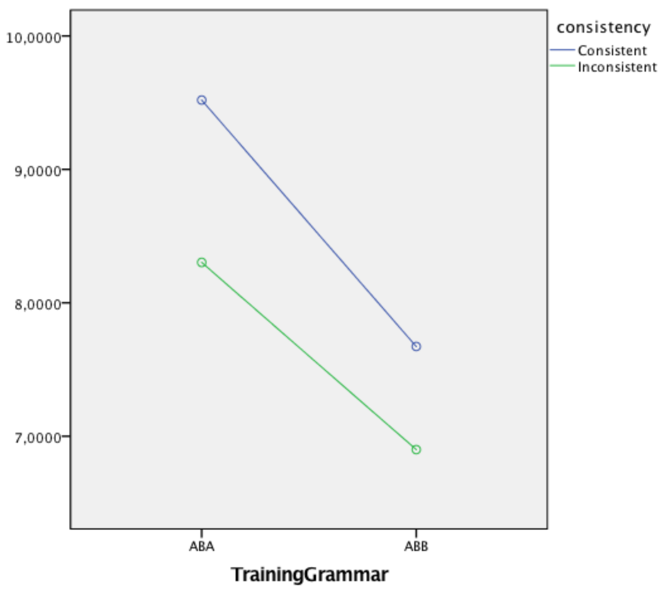

Figure 8. Looking time per consistency in the nonsense condition.

There was also a significant interaction between block, consistency and training grammar [F(2) =5,161, p = 0,011], which can be seen in figure 9 and figure 10. However, there is no significant interaction between training grammar and consistency [F(1) = 0,265, p = 0,613].

Bilingualism was analysed using a Repeated Measures ANOVA. The distribution of bilingual infants in the unknown condition can be seen in table 1. Bilingualism did not have a significant effect on the looking times [F(1) = 0,043, p = 0,838, bilingual (M = 7,857, SD = 1,505) monolingual (M = 8,260, SD = 1,229)].

Gender was also analyzed using a Repeated Measures ANOVA to see if this factor had a significant effect on looking times. The distribution can be seen in table 14. This factor did not have a significante effect on looking times [F(1) = 0,001 p = 0,972, female (M = 8,136, SD = 1,420), Male (M = 8,068, SD = 1,285)].

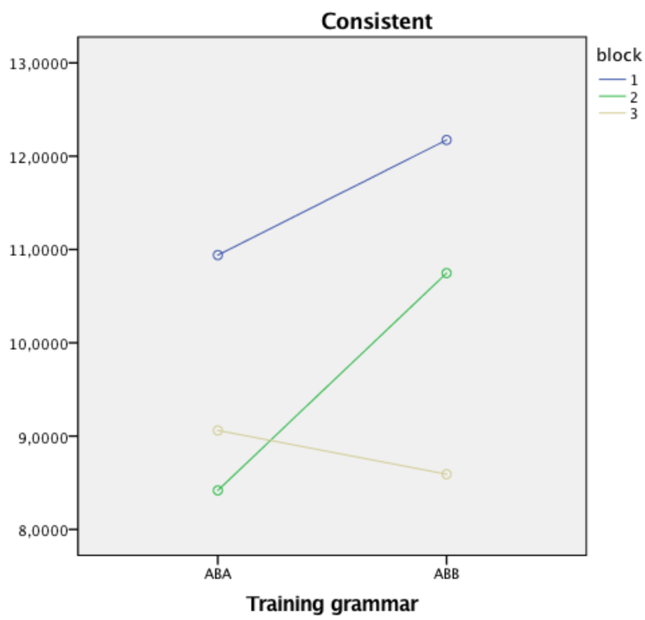

Figure 9. Looking time per training grammar per block for consistent

Figure 10. Looking time per training grammar per block for inconsistent

4. Discussion

The results show that our hypothesis was incorrect. We expected that the infants in would demonstrate a signifantly longer looking time to the inconsistent grammar in the known condition than the infants in the unknown condition. We also expected that the infants in the nonsense condition would not be able to differentiate between the grammars and would therefore show no significant difference in looking times. However, the only significant output in the known and unkown conditions are the decreasing looking times per block. This means that, as the experiment progressed, the infants looked less towards the screen. Against our expectations, the infants in the nonsense condition did show a significant difference in looking time between the consistent and inconsistent grammars, with longer looking times towards the consistent grammar. However, in line with the theory in Saffran et al (2007) and Marcus et al (1999), rule learning would be indicated with a significantly longer looking time towards the inconsistent grammar. However, Kidd, Piantadosi and Aslin (2010) demonstrated a “Goldilock effect”, which states that infants always seek to find an equilibrium for information absorption between overly predictable and overly unpredictable events. Therefore, shorter looking times can also indicate learning, as the infants no longer have to devote a lot of their attention towards events that are overly predictable. Thus, the results in the nonsense condition imply that the infants were able to differentiate between the consistent and inconsistent grammar, and that they show a familiarity preference towards the similar grammar.

Our aim was to research whether the conclusion of Saffran et al. (2007) – that visual rule learning is possible when the visual stimuli are familiar to the infants and easily categorizable – can be replicated using stimuli that, though familiar to the infants, are not part of the same category (e.g., images of only dogs). We hypothesized that lexical knowledge would imply sufficient familiarity and categorizability of the visual stimuli, and would enable the infants to successfully detect the same/difference relationship within every triad. Saffran et al (2007) argued that if the stimuli (e.g., abstract geometric shapes) are too abstract, the infant might treat every stimulus as separate from the rest, not as a member of the same category. In their experiment, infants who were reported to be highly interested in dogs performed best. Therefore, they argue, the presented stimuli should be created from materials that the infants are reported to be interested in. In the current experiment, we hypothesized familiarity is not binary; either known or unknown. Rather, we hypothesized that if the infant is very familiar with, and had sufficient lexical knowledge of, the objects used as stimuli then they would be able to generalize a grammatical rule. Objects that the infants might be familiar with, but of which they did not posses sufficient lexicical knowledge, were expected to negatively impact the infant's ability to generalize a grammatical rule. Therefore, the stimuli did not have to be domain-specific (i.e., part of the same lexicial category). However, as the N-CDI results show, the reported lexical knowledge of the infants for the visual stimuli in the known condition was lower compared to the *Lexilijst Nederlands*. This may have negatively effected the infants' ability to differentiate between the stimuli and to generate a grammatical rule.

We hypothesized that the infants in the known condition would be better at differentiating between the two grammars (i.e., showing a longer looking time towards the inconsistent grammar) than the infants in the unknown condition. We expected that the infants in the nonsense condition would not be able to differentiate between the two grammars due to uncategorizability and unfamiliarity of the stimuli and thus show no significant difference between looking time at the different grammars. However, as the results show, only block had a significant effect on looking times in the known and unknown condition. This means that looking time decreased per block. In these conditions, the infants also looked more towards the ABB grammar than to the ABA grammar, which may be caused by the duplication in the ABB grammar. Marcus (1999) and Saffran (2007) also found the duplication in the ABB grammar had an effect on looking times. To offset the possibility that the infants look longer towards the ABB grammars as a result of the duplication rather than learning, they conducted another experiment where they compared AAB to ABB. That infants are more susceptible to duplication in the testing grammar is also attested in Ota and Skarabela (2016), who argue that repetition-based rules are more easily learned, whether they are linguistic or non-linguistic in nature, than non-adjacent repetition or no repetition. Adult learners' performances in artificial grammar learning tasks also decrease when there are no repetition patterns in the test (Gomez, Gerken & Schvaneveldt, 2000; Tunney & Altmann, 2001). We expect that in a replication of this study using the grammars AAB and ABB, the effect of the preference for duplication is negated and will lead to more evenly distributed results gained from rule learning, without the effect of the preference for duplication. In the current experiment, infants in the known and unknown condition show a preference for the ABB grammar, which is in line with the argument posited by Ota and Skarabela (2016). Saffran et al. (2004) and Marcus et al. (1999) also tested this hypothesis, though the infants in their studies could also distinguish between ABA and ABB. To eliminate the possibility of duplication having an impact on looking time, an additional experiment will need to be run, containing the same conditions (known, unknown and nonsense) but with different grammars; ABB and AAB. This way, any influence duplication might have on the looking times of the infants, will be negated.

The results might also be caused by a lack of sufficient exposure during the familiarization phase. The familiarization phase in the current experiment was a replication of the familiarization phase in Saffran et al. (2007) and Marcus et al. (1999). We aimed to completely replicate the Saffran et al. experiment, but their description of the familiarization phase was scarce. Their description of the familiarization phase only mention that it was conducted via a habituation procedure. However, they do not mention the length of a trial, the maximum exposure, or whether it was infant-directed or experimentor-directed. Therefore, any changes that Saffran et al. made to the familiarization phase used by Marcus et al. (1999), were not mentioned. Therefore, we replicated the familiarization phase as described in Marcus et al. (1999). In our experiment the familiarization phase consisted of four blocks of 30 seconds of fixed exposure.

Between each familiarization trial there was a pause, during which the attention grabber played in order to center the infant's attention back on the screen. This delay might have had an impact on the infants' retention. Various researches have shown that a delay between 90 and 160 seconds can already have an effect on the retention of infants and adults (e.g. Rose *et al.*, 1981; Bahrick *et al.*, 1997; Bahrick & Pickens, 1995; Courage & Howe, 1998; Fagan, 1973; Richmond *et al.*, 2004). In a replication of this study, using dishabituation instead of familiarization might lead to better results, because then the learning phase is directed by the infant, rather than following a predetermined time period.

5. Conclusion

With this experiment we had expected to demonstrate that sufficient lexical knowledge is enough information for infants to successfully discrimate between two different grammars: the grammar they have been taught and a deviating grammar. Against our expectation, the infants did not demonstrate any learning in the known and unknown condition, but the results did indicate learning in the nonsense condition. The infants in the nonsense condition showed a significant difference in looking times between consistent and inconsistent, where they looked longer towards the consistent grammar. However, in this condition, the infants could not depend on their lexical knowledge to aid them in categorizing the stimuli and in generating a grammatical rule. That they look longer towards the consistent grammar than towards the inconsistent grammar, which we had expected based on the results in Saffran et al. (2007) and Marcus et al. (1999), may be explained by a familiarity preference or the Goldilocks effect (Kidd, Piantadosi, & Aslin, 2010).

We hypothesized that sufficient lexical knowledge would be enough information for infants to successfully generate a grammatical rule and to differentiate between grammatical and ungrammatical phrases. However, the infants in the known condition did not perform as we had expected. One explanation for their results are the reported N-CDI results (table 6). The infants appeared to have less lexical knowledge of the items used as the visual stimuli than we had anticipated, based on the numbers reported in the *Lexilijst Nederlands*. This might have made the experiment too difficult for the infants, as they lacked the necessary information to successfully detect a same/difference relationship in the stimuli and instead saw them as members of different categories. However, their familiarity with the items used as stimuli in the known condition, could be a reason for the difference in results between the known, unknown, and the nonsense conditions.

References

Bahrick, L., & Pickens, J. (1995). Infant memory for object motion across a period of three months: Implications for a four-phase attention function. *Journal of Experimental Child Psychology, 59*, 343–371

Bahrick, L., Hernandez-Reif, M., & Pickens, J. (1997). The effect of retrieval cues on visual preferences and memory in infancy: Evidence for a four-phase attention function. *Journal of Experimental Child Psychology*, 67, 1–20.

Benavides‐Varela, S., & Mehler, J. (2015). Verbal Positional Memory in 7-Month-Olds. *Child development, 86*(1), 209-223.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental psychology, 29*(4), 711.

Borensztajn, G., Zuidema, W., & Bod, R. (2009). Children's grammars grow more abstract with age—Evidence from an automatic procedure for identifying the productive units of language. *Topics in Cognitive Science, 1*(1), 175-188.

Courage, M., & Howe, M. (1998). The ebb and flow of infant attentional preference: Evidence for long-term recognition memory in 3-month-olds. *Journal of Experimental Child Psychology, 70,* 26–53.

Crain, S., & Lillo-Martin, D. (1999). *An Introduction to Linguistic Theory and Language Acquisition*. Malden, MA, USA: Blackwell

Fagan, J. (1973). Infants' delayed recognition memory and forgetting. *Journal of Experimental Child Psychology, 16*, 424–450.

Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109-135.

Gomez, R. L., Gerken, L., & Schvaneveldt, R. W. (2000). The basis of transfer in artificial grammar learning. *Memory & Cognition, 28*, 253–263.

Hauser, M. D., Weiss, D., & Marcus, G. (2002). Retracted: rule learning by cotton-top tamarins. *Cognition, 86*(1), B15-B22.

Horst, J. S. & Hout, M. C. (in press). The Novel Object and Unusual Name (NOUN) Database: a collection of novel images for use in experimental research. *Behavior Research Methods*. Manuscript accepted for publication.

Houston‐Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development, 13*(4), 341-348.

Johnson, S. P., Fernandas, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., & Slemmer, J. A. (2009). Abstract Rule Learning for Visual Sequences in 8- and 11-Month-Olds. *Infancy : The Official Journal of the International Society on Infant Studies, 14*(1), 2–18.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PloS one, 7*(5), e36399.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35-B42.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77-80.

Marcus, G.; Johnson, S.; Fernandes, K.; Slemmer, J. Rules, statistics and domain-specificity: Evidence from prelinguistic infants. Talk presented at the 29th Annual Meeting of the Boston University Conference on Language Development; Boston, MA. 2004.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101-B111.

Ota, M., & Skarabela, B. (2016). Reduplicated Words Are Easier to Learn. *Language Learning and Development,* 1-18.

Richmond, J., Sowerby, P., Colombo, M., & Hayne, H. (2004). The effect of familiarization time, retention interval, and context change on adults' performance in the visual paired‐comparison task. *Developmental Psychobiology, 44*(2), 146-155.

Roder, B. J., Bushnell, E. W., & Sasseville, A. M. (2000). Infants' preferences for familiarity and novelty during the course of visual processing. *Infancy, 1*(4), 491-507.

Rose, S. A. (1981). Developmental changes in infants' retention of visual stimuli. *Child development*, 227-233

Saffran, J.R., Aslin, R.N., Newport, E.L., (1996).Statistical learning by 8-month-old infants. *Science*, (274)1926–1928.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27-52.

Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language, 47*(1), 172-196.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*(3), 669-680.

Schlichting, J.E.P.T., and Lutje Spelberg, H.C. (2002). *Lexilijst Nederlands: Een Instrument om de Taalontwikkeling te Onderzoeken bij Nederlandstalige Kinderen*

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category – ELAN and ISO DCR. *In: Proceedings of the 6th International Conference on Language Resources and Evaluation* (LREC 2008).

Tunney, R. J., & Altmann, G. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 614–639.

Valian, V. (1986). Syntactic Categories in the Speech of Young Children. *Developmental Psychology, 22*. 562-579.

Zink, I., & Lejaegere, M. (2002). N-CDIs: *Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDIs van Fenson et al. Leuven/Leusden: Acco*