Pauses matter: Rule-learning in Children

Anika van der Klis*
Utrecht University, the Netherlands

Rianne van Lieburg*
University of Antwerp, Belgium

Lisa Lai-Shen Cheng
Clara Cecilia Levelt
Leiden University, the Netherlands

* these authors contributed equally to this work

Abstract: Language learners have to both segment words and discover grammatical rules connecting those words in sentences. In adult listeners, the presence of a prosodic cue in the speech stream, for example, a pause, appears to facilitate rule-learning of non-adjacent dependencies of the form A_iXC_i (Peña et al., 2002). Only when listening to the artificial language containing pauses, could participants identify rule-words of the form $A_iA_jC_i$ or $A_iC_jC_i$, where intervening syllables were moved from A- or C-positions. Frost and Monaghan (2016) found in a similar study that participants who were tested with novel, rather than moved, intervening syllables in A_iXC_i items showed rule-learning even when the familiarisation stream contained no pauses. The present study re-examines the facilitative effect of pauses in discovering structural rules in speech in a novel population: children aged 7-11. We used the same artificial speech stimuli as Peña et al. (2002) and tested children in both a moved-syllable and novel-syllable forced-choice task. The results of 140 children show that pauses provide a facilitative effect on rule-learning – also for young learners. Regardless of syllable types, only children who listened to the familiarisation stream containing pauses chose words following the rule above chance-level.

Keywords: artificial grammar learning; statistical learning; non-adjacent dependencies; prosody; school-age children

Corresponding author(s): Rianne van Lieburg, Centre for Computational Linguistics and Psycholinguistics, University of Antwerp, Grote Kauwenberg 18, 2000 Antwerpen, Belgium. Email: rianne.vanlieburg@uantwerpen.be

ORCID ID(s): https://orcid.org/0000-0001-5024-8096

Citation: van der Klis, A., van Lieburg, R., Cheng, L.L., & Levelt, C.C. (2022). Pauses matter: Rule-learning in children. *Language Development Research*, *3*(1), 44–64. http://doi.org/10.34842/2023.0466

Introduction

Language learners need to both segment words and discover grammatical rules connecting those words in sentences. Saffran et al. (1996a) demonstrated that 8-month-olds were able to segment words from running speech after a short exposure based on statistical relationships between neighbouring speech syllables. They could infer word boundaries between two syllables with a low transitional probability in the sequence (i.e., a transitional probability of 0.33 between words versus a probability of 1.0 within words). Adults have also been shown to use dips in transitional probability to infer word boundaries and to successfully extract words from a continuous speech stream (e.g., Saffran et al., 1996b; Peña et al., 2002).

Peña et al. (2002) suggested that, while statistical relationships are sufficient for speech segmentation, additional cues are needed for the detection of grammatical rules. The artificial language to which their adult participants were exposed consisted of trisyllabic sequences that followed a non-adjacent dependency (NAD) rule of the form A_iXC_i, where A_i always predicted C_i (e.g., the English progressive *is X-ing*). Hence, the transitional probability between A_i and C_i was 1.0. The within-word transitional probability (between A_i and the adjacent X or between X and the adjacent C_i) was 0.33 while between words (between the last syllable of any item and the first syllable of the next item) it was 0.5. Peña et al. (2002) showed that, when presented with a continuous speech stream, listeners only deemed test items that had appeared in the same form in the stream correct (i.e., they could segment the trisyllabic items) but deemed test items constructed by moving an A or C syllable to the X-position, resulting in $A_iA_jC_i$ or $A_iC_jC_i$, incorrect (i.e., they could not find words following the rule when there was an intervening element originating from another position in the stream). When gaps of 25-ms, which Peña et al. (2002) called “subliminal pauses”, were placed between each A_iXC_i triplet (e.g., puRAki-pause-puLIki) in the familiarisation stream, i.e. segmenting the stream into smaller constituents, adults *were* able to extract possible words that followed the NAD rule (hereafter rule-words) containing moved syllables. This showed that participants could identify the NAD rule when – and only when – pauses were added to the otherwise identical familiarisation stream.

Peña et al. (2002) assumed that streams that are segmented by pauses relieve listeners of the task of computing probabilities to segment words, thereby giving them the chance to discover underlying rules. This hypothesis is in line with studies suggesting that prosodic cues that mark the boundaries of constituents may have a scaffolding function during language acquisition (e.g., Soderstrom et al., 2003; Morgan, 1986). Several studies showed that adults are not able to extract and generalise NADs from continuous speech streams without perceptual cues marking phrases. Similar to Peña et al. (2002), Endress and Bonatti (2007) showed that adults only preferred class-words (of the type A_iXC_j) to part-words (of the type XCA or CAX) when listening to a familiarisation stream containing 25-ms pauses. Without these cues, participants did not show a preference for either class-words or part-words. Marchetto and Bonatti (2013, 2015) also examined children (12-month-olds and 18-month-olds) using head-turn procedures and found evidence that they could learn NADs of the type A_iXC_i, but only when listening to a stream containing either 25-ms or 200-ms pauses. The authors proposed that the same learning mechanism used by adults might be readily available for infants - triggered by the same acoustic properties in the stream. Grama et al. (2016) examined whether other types of perceptual cues affect the learning of artificial NADs in adults. They found that performance in a forced-choice task increased when the dependent elements (i.e., A_i and C_i in A_iXC_i strings) were either acoustically enhanced or reduced in the familiarisation stream, but only when the A_iXC_i strings were also separated by pauses. This led to the hypothesis that NAD learning is easiest when the dependent elements are both perceptually distinctive *and* integrated into a prosodically natural contour (Grama et al., 2016). These behavioural results show that NAD learning in both infants and adults is enhanced when prosodic cues are present that break up the continuous speech stream into smaller constituents containing the rule (i.e., A_iXC_i strings). These smaller constituents may play a facilitative role to learners extract rules.

The processing of NADs in the brain has been studied too, using electroencephalography (EEG). Mueller et al. (2008) found that adult participants showed different event-related potential (ERP) patterns when listening to a stream containing pauses, in addition to an increase in correct responses by 30% in a condition with pauses compared to a condition without pauses. In the condition with pauses, participants showed an additional positivity in their responses, which the authors interpreted as reflecting more controlled, attention-guided mechanisms. De Diego-Balaguer et al. (2015) also examined ERPs in adults while they listened to different artificial speech streams containing trisyllabic items with and without 25-ms pauses in between them. Their results showed that pauses altered electrophysiological responses to the stream. In the stream without pauses, the amplitude N1 component increased at syllable onsets, which indicates that participants pay attention to them for the sake of locating word boundaries. Pauses reduced the mean amplitude of the N1 component in the first syllable of the trisyllabic items, which may indicate that participants segment the stream by means of the pauses, and no longer need to orient to the syllable onset for the location of the word boundaries. Behavioural results of this study also showed that while participants were indeed better at segmenting words when the continuous speech stream contained pauses, these pauses did not improve rule learning (de Diego-Balaguer et al., 2015). The findings of these studies corroborate Peña et al.'s (2002) hypothesis that the availability of perceptual cues relieve listeners of the speech segmentation task and alter processing of the speech stream, but not their necessity for rule-learning.

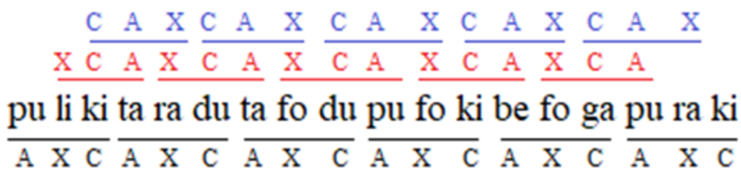

Figure 1. ***A short excerpt of the familiarisation stream used by Peña et al. (2002) containing six AXC rule-words (black) and ten part-words of the type XCA (red) and CAX (blue).***

Nonetheless, the interpretation that statistical relationships alone do not suffice for rule-learning is heavily debated in the literature (see Perruchet et al., 2004; Endress & Mehler, 2009; Frost & Monaghan, 2016). Peña et al. (2002) concluded that participants did not learn the rule in the non-pause condition because participants significantly chose part-words (i.e., of the form XCA or CAX) over rule-words (i.e., of the form $A_iA_jC_i$ or $A_iC_jC_i$). However, Frost and Monaghan (2016) pointed out that, even though infrequently, part-words had appeared in the familiarisation stream exactly as such, while rule-words containing moved syllables, such as *pubeki*, had not. The artificial language stream consists of many adjacent words, and part-words were formed by syllables that span word boundaries, as illustrated in Figure 1. This excerpt of only six rule-words in fact contains ten part-words. Participants were thus forced in the test to choose between part-words, that had appeared in the familiarisation stream, and rule-words with moved syllables, that had never occurred as such. Preferring rule-words over part-words, then, requires not only identification of the structural generalisation, but also suppression of learned (or encountered) syllable sequences. Frost and Monaghan (2016) used the same artificial language used by Peña et al. (2002) in their study and created test items where the intervening elements were either moved (from A or C positions) or completely novel. Their 10.5-min long familiarisation stream did not contain pauses or any other prosodic cues. Adults selected rule-words rather than part-words significantly above chance, but only when the test items contained novel, rather than moved, intervening elements ($M = .693$, $p < .001$). They therefore concluded, in line with Endress and Mehler (2009) and Perruchet et al. (2004), that the pause used by Peña et al. (2002) only served as an additional cue, increasing the saliency of the positions of individual syllables, rather than relieving listeners from the segmentation task. The result from Frost and Monaghan (2016) further questions the actual role of prosodic cues that mark constituent boundaries in rule-learning. The experiments reported in the present paper aim to inform this debate.

Currently, it is not known whether school-aged children show performance similar to adults, because this is an underrepresented age group in the rule-learning literature. Research has yet to find out whether school-aged children, like adults, can generalise NAD-rules over test items with novel syllables. In Soderstrom et al. (2007), 16-month-old infants could only generalise NADs in a natural language to novel nonsense stems (e.g., vod teebs) if they were first presented with familiar stems (e.g., dog runs). The authors hypothesised that infants had been distracted by the presence of unfamiliar words in the stimuli. Similarly, Grama and Wijnen (2018) showed, using an artificial language paradigm, that while 18-month-olds do have abstract knowledge of A_iXC_i strings after exposure, they cannot generalise the NADs when there are novel intervening syllables. Novel items may actually draw children's attention away from the dependency, yielding hindering, rather than facilitating effects on rule learning. These results are in contrast with the findings by Frost and Monaghan (2016) for adults, for whom the use of novel X-syllables did not hinder the ability to generalise the NADs. The present study is the first to assess artificial rule-learning using novel stimuli in school-age children.

It is also not known if school-age children can successfully learn NAD-rules during passive listening. Mueller et al. (2012) found that adults were less successful than infants in NAD learning under passive conditions when measuring ERPs. In an ERP experiment with 2- and 4-year-olds, Mueller et al. (2019) showed that passive learning of NADs, in an artificial language with pauses, declined between 2 and 4 years of age. This is linked to maturation of the Prefrontal Cortex (PFC), which is completed around the age of 7 and involves a switch to a different, more adult-like learning mechanism (Skeide & Friederici, 2016). Similarly, van der Kant et al. (2020), using functional near-infrared spectroscopy (fNIRS), found evidence that only 2-year-olds, but not 3-year-olds could detect linguistic NAD violations during passive listening. In a recent study using ERPs, Paul et al. (2021) examined children between 1 and 3 years old, and although all children showed evidence of learning NADs in a foreign language, there was a gradual decrease in the strength of this evidence across these ages. It may therefore be possible that during development, there is a decline in the ability to learn through passive listening when there are no additional cues that mark the NADs. This suggests that children aged 7 to 11 may not be as successful as infants in detecting NADs during passive listening.

Previous studies have shown that adults outperform children when NAD learning is assessed using a task that requires more declarative knowledge, such as a grammatical judgement task, even when they do not receive instructions prior to listening to the speech stream. Ferman and Karni (2010) examined artificial grammar learning in adults, 12-year-olds, and 8-year-olds. Adults outperformed both groups of children, and 12-year-olds outperformed 8-year-olds. This was reflected in higher accuracy as well as shorter response times in both a grammatical judgement and a production task. Ojima and Okanoya (2020) tested adults and children aged 5 to 12 on centre-embedding learning in an artificial A^nB^n grammar which also generates NADs.

They found that the majority of the adults could generalise the rules to novel stimuli in a go/no-go task, indicating that they had learned the grammar. However, only about a quarter of the children in their study succeeded in this. The authors suggested that failure in this task by all the other children is due to memory constraints, not due to rule-learning deficits. Lammertink et al. (2019) examined NAD learning in children aged 5 to 8 (*M* = 7.3) years old. They did find evidence for sensitivity to NADs in online reaction times, but above-chance performance was not found in an offline forced-choice task. The children in their study did not receive explicit instructions, but their task did require a certain level of attention to the stimuli. The authors argued that this grammatical judgement task required more metalinguistic awareness and attention, which is more difficult for children compared to adults. Marimom et al. (2021) ran a similar experiment with children aged 3 to 8 (*M* = 6.2) years old, but they used a stem completion task instead of a forced-choice task. They found evidence of learning in reaction times, and above-chance performance at the group level during the stem completion task. The results furthermore showed faster reaction times for older children, although their accuracy scores did not increase. Importantly, Marimom et al. (2021) added 20-ms pauses at the beginning of each AXB stimulus as well as a longer pause between the A-element and X-element, which may have enhanced children's performance in this study. Schaadt et al. (2020), in a study with 7-year-olds, found no significant above-chance performance at the group level on a grammatical judgement task (choosing between *correct* or *incorrect*), after familiarisation with short sentences - separated by pauses - containing NADs in a natural foreign language. However, accompanying ERP data did show, both at the group and at the individual level, that especially after a retention period of one night's sleep, a representation of the NAD had been built. Children can implicitly recall NADs, as shown by more negative ERP responses to NAD violations, but they do not show explicit knowledge in the grammatical judgement task. The authors concluded that their grammatical judgement task was still too difficult for children of this age, and that they might have been able to show an effect of learning in a forced-choice task. The present study uses this more suitable task and aims to add to our understanding of NAD learning in this age group.

In our study, we investigated the performance of 7- to 11-year-old children who were tested on both moved and novel intervening syllables in the AXB test items. We expected children above the age of 7 to have switched to a more adult-like mechanism (Skeide & Friederici, 2016) that benefits both from additional segmentation cues in the speech stream (e.g., Peña et al., 2002; Grama & Wijnen, 2018) and from a task which requires more metalinguistic knowledge to guide their attention to the NADs (e.g., Pacton & Perruchet, 2008; Bialystok, 1986) compared to younger children. In the first experiment, we used the same A_iXC_i language as in Peña et al. (2002) and created test items using moved syllables of the form $A_iA_jC_i$ or $A_iC_jC_i$ as intervening syllables. In line with Peña et al. (2002), we expect better learning of the NADs when pauses were present in the familiarisation stream. Pauses segment the stream into constituents, which draws more attention to the dependent elements on constituent boundaries. This could help children discover rules. However, this experiment does not specifically address the question of whether a segmented stream also facilitates the discovery of NAD-rules when participants are presented with novel intervening syllables, or whether using novel elements alone is sufficient to draw children's attention to underlying rules. In our second experiment, we tested a new group of children and used novel syllables as intervening syllables in the test items. Here there are two possible outcomes; either the novel intervening syllables are sufficient for drawing children's attention to the underlying rule without needing other segmentation cues, as Frost and Monaghan (2016) found for adults, or the novel intervening syllables end up hampering children's ability to generalise, as was found for infants (e.g., Soderstrom et al. 2007; Grama and Wijnen, 2018). If the presence of pauses results in more successful learning in the novel-syllable task, this constitutes more precise evidence for the facilitative effect of pauses. By pitting moved syllables against novel syllables, we can get a deeper understanding of the effect of pauses in artificial rule-learning in school-age children.

Method

Participants

For the first experiment, we aimed to collect as many data as we could within the duration of one semester[1]. We tested 92 children (55 boys, 37 girls[2]) between 7- and 11-years-old (*M* = 8.55, *SD* = 1.18). For the second experiment, we tested a new group, again within the duration of one semester, collecting data from 51 children (27 boys, 24 girls) aged between 7- and 11-years-old (*M* = 9.04, *SD* = 1.06). We excluded three participants because they did not follow the test instructions properly.[3] All children were native speakers of Dutch and did not report any hearing or language-related problems. They were recruited from different primary schools in the Netherlands (Leiden and Rotterdam area) and the Leiden University Babylab participant database. The parents gave their written consent and filled out a short questionnaire providing information concerning the inclusion criteria. After participating, all children

[1] The study was funded by and conducted within the Research Traineeship Programme at Leiden University, which had a fixed duration of one year.

[2] In Experiment 1, there is an imbalance in gender ratio: we tested more boys than girls. Exploratory analyses did not reveal any effect of gender in either Experiment 1 or Experiment 2.

[3] We excluded two participants from Experiment 1: one in the condition with pause (18 correct responses) and one in the condition without pauses (19 correct responses). One participant in the without-pauses condition was excluded from Experiment 2 (14 correct responses).

received a monetary compensation of five euros.

Previous studies reported strong effect sizes with adults: Frost and Monaghan (2016) report a Cohen's d of 1.2 on the novel syllable task testing 18 participants. Because children usually show much more variability in performance, we used two-thirds of this factor size for our power estimate: $d = (2/3 * 1.2) = 0.8$, which corresponds to an odds ratio of 4.3.[4] We used the WebPower package (Zhang & Yuan, 2018) to determine the minimum sample size required to have at least 80% power. Results indicated that the minimum sample size should be $n \geq 43$. In both our experiments our sample sizes exceeded this number.

Materials

The familiarisation stream consisted of a ten-minute long sequence of syllables created with the "Female 5" French voice[5] of the speech synthesiser Praat (Boersma & Weenink, 2018). We used the same A_iXC_i language and the same syllables as in Peña et al. (2002). The A_iXC_i dependencies were *pu*X*ki*, *ta*X*du* and *be*X*ga*. The X-syllables were *li*, *ra* and *fo,* leading to 9 different trisyllabic items: *puliki, puraki, pufoki, talidu, taradu, tafodu, beliga, beraga,* and *befoga.* It should be noted that the A_i and C_i syllables involved plosives, and the X syllables continuants, which resembles natural language (cf. Frost and Monaghan, 2016). We pseudorandomized the order of the different A_iXC_i sequences ("words") according to the same criteria as in Peña et al. (2002). Each trisyllabic item occurred a hundred times in the stream. Two subsequent items never started with the same syllable. The X-syllable always differed between two subsequent items. The transitional probability between A_i and C_i was 1.0, the within-word transitional probability (between A_i and the adjacent X or X and the adjacent C_i) was 0.33 and the between-words transitional probability (between the last syllable of any item and the first syllable of the next item) was 0.5. We created two versions of the familiarisation stream: one containing a 10-ms[6] pause between each trisyllabic item and one without such pauses, leaving them completely identical otherwise. We

[4] Odds ratio = $e^{d \times \frac{\pi}{\sqrt{3}}}$ (see Sánchez-Meca et al. 2003).

[5] Peña et al. also used a French voice, but note that their participants were native speakers of French. We also used a French synthesiser rather than a Dutch synthesiser, because the phoneme /g/ in the used syllable /ga/ is not available in a Dutch synthesiser as Dutch only has /g/ in loanwords (meaning that the children were still familiar with /g/). Crucially, we told the children that they were going to listen to a foreign language.

[6] Peña et al. (2002) reported the use of 25-ms pauses in the familiarisation stream. Since different speech synthesisers may treat such settings differently, we measured the actual pause duration between trisyllabic items from Peña et al. using Praat and mimicked those pause durations with the speech synthesiser we used. With our speech synthesiser, generating a 10-ms pause resulted in a familiarisation stream with pauses that were comparable to the ones used by Peña et al.

used 5-second fade-in and fade-out effects, following Peña et al. (2002) and Frost and Monaghan (2016), to ensure that there was no audible first or last syllable.

In the first experiment, the forced-choice task included 36 pairs of rule-words following the A_i _C_i rule with moved intervening syllables originating from A or C positions filling the _ slot (e.g., *pubeki of the type AAC*) and part-words of the types CAX (e.g., *gapufo*) or XCA (e.g., *fogapu*). Audio files of the rule-words and part-words were created in Praat using the same synthetic voice as the familiarisation stream. In the second experiment, the forced-choice task was adapted by creating test items with novel intervening syllables (i.e., *ve*, *no* and *si*) that had never occurred in the familiarisation stream. Both the rule-word and the part-word contained these novel syllables instead of moved syllables. The novel syllables contained continuants, like in the X-syllables in the familiarisation stream. We used different novel syllables from Frost and Monaghan (2016) (who used *ve, zo* and *thi*), in order to only use phonemes with which the children are familiar from Dutch. The forced-choice task was both programmed and run with Praat. The script containing the forced-choice task is provided in the supplementary materials.

Procedure

The experimental procedure in both experiments consisted of a familiarisation phase followed by a test phase. First, the Dutch children listened to a short excerpt from a British English television show, and they were asked whether they recognised the language, to which the majority responded that they recognised it as being English - or at least as a foreign language they heard before. We used this to explain that they were going to listen to another language, but one that they had never heard before. We instructed them that we were going to see whether they could also recognise this new language afterwards. We did not explicitly explain that they should look for rules. The children watched a video of *Pingu* (a children's animation show) while they were presented with the familiarisation stream through over-ear headphones. Participants randomly received the familiarisation stream either with or without 10-ms pauses between the A_iXC_i items.

After the familiarisation period, the children were presented with the forced-choice task. The test started with three practice items that were not included in the analysis. All children received the same pairs of test items in random order. Participants were asked to choose the item which most likely belonged to the familiarisation language (instructed as "which one belongs to the language you just heard?"). After listening to a pair consisting of a rule-word and a part-word, two big buttons with "1" and "2" appeared on the screen. The children were instructed to select either "1" or "2" using the computer mouse to select the first or second word of the pair. They could listen to the pair one more time by clicking on a replay button. The majority of participants immediately started selecting words from the word pairs, and we did not provide them with any additional instructions. When children were reluctant to answer during the practice phase, we reassured them that they could go with their first intuition, and that they did not have to be certain about their answer. The testing phase took about ten minutes for each child to complete. This experiment was run using Praat on Windows computers.

Coding and analysis

The responses to all test pairs were automatically coded as "correct" (i.e., the participant selected the rule-word) or "incorrect" (i.e., the participant selected the part-word) by the Praat script. This resulted in a list of 36 answers, correct or incorrect, for each participant, which we used as the binary outcome variable in a generalised linear mixed-effects model with the presence of pauses as a fixed effect predictor. Each participant also received a final score between 0 and 36 correct answers. We used this to examine whether participants scored at an above chance-level (i.e., showed a learning effect). In addition, we calculated which scores were outliers. An outlier was defined as being three times the *SD* above or below the mean. Outliers were excluded from the analyses.

We analysed the results of the two experiments separately to facilitate the comparison of the results of the first experiment to the results of Peña et al. (2002), and those of the second experiment to the results of Frost and Monaghan (2016). In addition, we performed a joint analysis to further assess the relationship between the use of moved or novel syllables and pauses in NAD learning.

Results

Separate analyses

In the first experiment using test items constructed with moved intervening syllables, we analysed the results of 91 children. Children who listened to the stream without pauses (n = 46) had an average score of 16.31 correct responses (45.6%, *SD*: 3.96), whereas those who received the stream with pauses (n = 45) had an average score of 19.20 (53.3%, *SD*: 4.47) correct responses.

The responses were compared between the two groups by fitting a binomial generalised linear mixed model (R-package lme4, Bates et al., 2012). The presence of pauses in the familiarisation stream and children's age in months were included as fixed effects. Gender or an interaction between Pause and Age did not improve model fit so we report the model without them. We centred and scaled Age to increase convergeability. A random intercept was included for participants which significantly improved model fit. The *p*-values were obtained by using the package lmerTest (Kuznetsova et al., 2017). We used the jtools package (Long, 2020) to calculate exponentiated coefficients (i.e., odds ratios).

The significant negative intercept indicates that the responses of participants who were exposed to the familiarisation stream without pauses were more often incorrect than correct. Significantly more items were answered correctly if the participant had received the familiarisation stream with pauses ($p < .01$). In addition, scores improved upon increasing age ($p < .05$). When test items contained moved intervening syllables, children who received the stream with pauses were 1.32 times more likely to give a correct answer than children who received the stream without pauses. For both conditions, we also compared the proportion of correct responses to chance-level (50%, which equals a mean of 18 correct responses) in a one-tailed *z*-test[7]. The group that received the familiarisation stream without pauses did not perform above chance-level ($p = .99$), whereas the group that listened to the familiarisation stream with pauses did ($p < .01$, $d = 0.268$).

Table 1. ***Results of the generalised linear mixed model of the first experiment using moved X-syllables (n = 3276, log-likelihood = -2245.0)***

Predictor	Estimate	Exponent. Estimate	*SE*	Wald *Z*	*p*-value
(Intercept)	-0.14	0.87	0.07	-2.00	0.05
Pause	0.27	1.32	0.10	2.74	0.01
Age	0.11	1.11	0.05	2.18	0.03

In the second experiment, using test items constructed with novel intervening syllables, the results of 49 participants were analysed. Children who listened to the familiarisation stream without pauses ($n = 24$) had an average of 18.25 (50.7%, *SD*: 3.88) correct responses, whereas children who received the familiarisation stream with pauses ($n = 25$) had an average of 19.62 (54.5%, *SD*: 4.59) correct responses. One participant, who was exposed to the familarisation stream with pauses, had an outlier score of 33 correct responses. After excluding this participant[8], the group of children

[7] We chose to perform a one-tailed z-test, because we tested for a positive learning effect. We report on the two-tailed z-tests of all our statistical comparisons against chance level in the Supplementary materials. The levels of significance remain the same, except that the group in the moved-syllables experiment that received the familiarisation stream without pauses scored significantly below chance-level ($p < .001$).

[8] The statistical analyses including the outlier are reported in the Supplementary materials. Our conclusions are not altered by this inclusion.

who received the familiarisation stream had an average of 19.20 (53.6%, *SD*: 3.76) correct responses.

We compared the responses of the two groups by fitting a binomial generalised linear mixed model (R-package lme4, Bates et al., 2012). Neither the presence of Pause (= 0.95, p = .33) nor Age (centred and scaled) (= 0.78, p = .38) nor their interaction (χ = 0.57, p = .45) improved model fit when using novel intervening syllables in the test items. The model was only improved by adding a random intercept for participants to the null model. The fixed factors do not contribute beyond the random effect to explain differences in the number of correct responses. Again, we also compared the proportion of correct responses to chance-level per condition in a one-tailed *z*-test. The group that was exposed to the familiarisation stream without pauses did not perform above chance-level (p = .35), whereas the group that was presented with the familiarisation stream with pauses did (p = .001, d = 0.319).

Joint analysis

To examine a possible interaction between the moved- and novel- syllable conditions, we further analysed the results by performing a joint analysis of both experiments (n = 139). We built up the model by adding a random intercept for participants which significantly improved model fit. Then, we added fixed effects step by step. We found significant improvements of fit when adding Pause and Age. An interaction between Pause and Age did not significantly improve model fit. Neither Gender nor Syllable Type improved fit. We report the final model with fixed effects of Pause and Age and a random intercept for participants in Table 2.

The intercept represents the log-odds for a correct response in the condition of exposure to the familiarisation stream without pauses and the forced-choice task of test items with moved intervening syllables. Significantly more items were answered correctly if the participant had received the familiarisation stream with pauses (p < .01). In addition, there was a positive effect of age (p < .01). Across both experiments, children who listened to a stream with pauses were 1.25 times more likely to give a correct answer compared to children who listened to the stream without pauses, regardless of syllable type.

Table 2. ***Results of the generalised linear mixed model of the joint analysis (n = 5004, log-likelihood = -3438.7)***

Predictor	Estimate	Exponent. Estimate	*SE*	Wald *Z*	*p*-value
(Intercept)	-0.10	0.91	0.05	-1.76	0.08
Pause	0.22	1.25	0.08	2.88	0.007
Age	0.11	1.11	0.04	2.71	0.004

In our model, we compared the results of the different groups but not the learning effect per se, i.e., scoring higher than chance-level. In the separate analyses of both groups, we found above chance performance in the condition with pauses, but not in the condition without pauses. Overall, participants did not score above chance (50%) (M = 18.13, SD = 4.23, n = 139, p = .31 in a one-tailed z-test). However, like in the analyses of the individual experiments, a one-tailed z-test showed that the group that listened to the familiarisation stream without pauses did not perform above chance (M = 17.04, SD = 4.02, n = 70, p > .99), whereas the group exposed to the familiarisation stream with pauses did (M = 19.23, SD = 4.18, n = 69, p = <.001, d = 0.294).

The one-tailed z-test uses mean scores. However, we also looked at individual scores (see Figure 2). For an individual score significantly higher than chance-level, at least 24 out of 36 items should be correct. Note that 'significantly higher than chance-level' means a score higher than 95% of the scores in a binomial distribution with a probability of success of 50%. The probability of a score of 24+ correct responses is 3.26%[9]. Under a binomial distribution, we would therefore expect 5 out of 140 participants to have 24+ correct items. However, 15 participants (11.4%) turned out to have 24+ correct responses, 12 of which received the familiarisation stream with pauses. Of the 3 participants who were exposed to the familiarisation stream without pauses, 2 were in the moved intervening syllable condition (4.3%) and 1 was in the novel intervening syllable condition (4.2%). In the conditions with pauses, many more participants than expected scored above chance-level (X^2 (1, N = 69) = 8.49, p < .01)). In the conditions without pauses, the number of participants that scored above chance-level was not higher than expected (X^2 (1, N = 140) = 1, p = 1)).

The experiments in this study investigated NAD learning abilities in children aged 7-11 in an artificial A_iXC_i grammar. Our study aimed to assess whether perceptual cues in the speech stream facilitate the learning of NADs in children. We examined this by pitting moved syllables against novel syllables as intervening elements in A_iXC_i strings to get a deeper understanding of the effect of pauses on NAD detection. We found that the presence of pauses indeed facilitated the detection of NADs in these young learners, regardless of syllable type in the test phase, though pauses did not guarantee that all children discovered the rule. Children who listened to the familiarisation stream with pauses chose rule-words significantly more often than part-words in a forced-choice task. When testing children using moved intervening syllables, we found a large improvement when pauses were added to the stream. The

[9] We calculated the probability of individual scores using the probability density function of the binomial distribution: $P(p,n,r) = p^r \times (1-p)^{n-r} \times \frac{n!}{r! \times (n-r)!}$, where P is the probability of a particular outcome, p is the probability of success of each trial, n is the number of trials, and r is the number of successes. We calculated at which score the cumulative probability was less than 0.05.

percentage of chosen rule-words in the forced-choice task showed a significant increase after familiarisation with the speech stream including pauses, replicating the findings by Peña et al. (2002) in a novel population: school-age children.

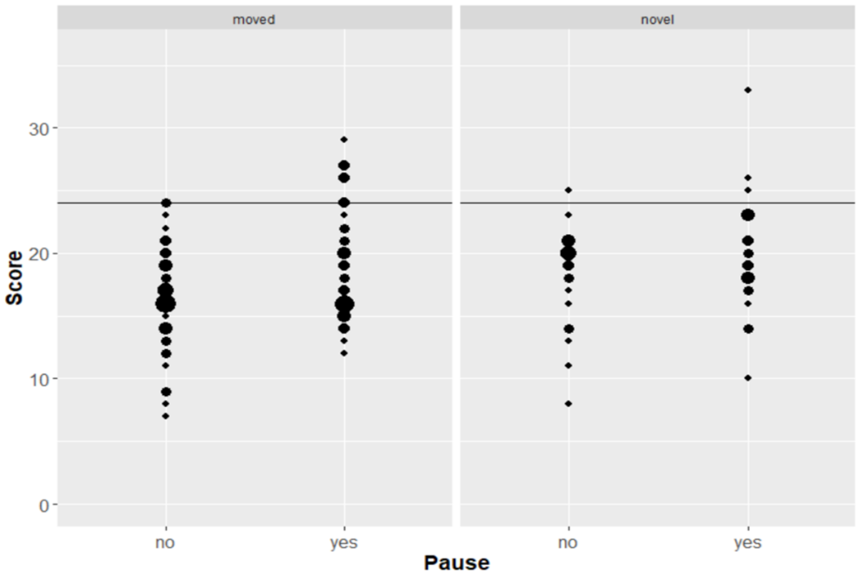

Figure 2. ***Bubble chart of number of the correct responses per condition, reference line at 24 correct responses.***

General Discussion

With regard to the effect of pauses when testing novel intervening syllables compared to the effect in the experiment with moved intervening syllables, the results are not as straightforward. There is a discrepancy between the outcomes of the separate analysis and the joint analysis. In the separate analysis of Experiment 2, there is no significant effect of Pause on scores. On the other hand, the one-tailed z-test showed that participants only performed significantly above chance-level (i.e., a score of 50%) when the familiarisation stream contained pauses. In addition, the joint analysis revealed no interaction between Syllable Type and Pause and a significant effect of Pause. Pauses facilitated the learning of dependency relations in both moved-syllable and novel-syllable conditions when analysing the combined data. We suggest that the novel intervening syllables in the test items led to too much variability in performance to detect any effect when analysing this dataset alone. Another possibility is that the difference between the condition with pauses and the condition without pauses is larger in the experiment with moved intervening syllables than in the experiment with novel intervening syllables, due to the below-chance performance found in the moved-syllable condition after familiarisation without pauses. Interestingly, in the novel-syllable condition, we do not find the same below-chance performance. The enhanced performance in Experiment 2 strengthens the idea that the part-words encountered in the familiarisation stream in Experiment 1 were harder to reject compared to rule-words of the form A_iAC_i or A_iCC_i because part-words were statistically more likely in their original form, and not because children were not able to generalise the NADs. When listening to a continuous stream without pauses, this resulted in below-chance performance in Experiment 1 (45.6%), which disappeared in Experiment 2, with overall results remaining at chance-level (50.78%).

Using novel intervening syllables in the test stimuli may inhibit participants' need to suppress any part-words that had occurred in the stream spanning word boundaries (as in Experiment 1), but it does not enhance learning in such a way that children no longer need other segmentation cues. This result suggests that while using novel intervening syllables in the test stimuli does prevent below-chance performance, due to the inability to reject part-words that occur in the familiarisation stream, it does not allow children to detect the underlying NADs. This is in contrast with the results found by Frost and Monaghan (2016). In their study with adults, using novel intervening syllables in the test stimuli yielded evidence for learning without any other cues segmenting the familiarisation stream.

In the current study, children only chose significantly more rule-words than part-words containing novel intervening syllables when the familiarisation stream was segmented by pauses. When the familiarisation stream contained pauses, we observed an increase to 53.6% correct scores in Experiment 2. This is significantly above chance-level (i.e., a score of 50%). In other words, school-age children do not benefit from novel intervening syllables in the way that adults do. In contrast to the findings by Soderstrom et al. (2007) for infants, we did not find a clear hindering effect of novel elements either, although children did seem to benefit less from the presence of pauses when being tested on novel intervening syllables. In the mixed-effects model, Pause did not significantly contribute beyond the random effects to explain differences in performance when testing children using novel intervening syllables. However, when comparing the proportion of correct responses against chance-level, we found that only the group of children who received the familiarisation stream with pauses performed above chance. These results suggest that school-age children are more likely to extract an artificial NAD rule during passive listening when an additional segmentation cue in the form of a pause is present in the speech stream, and they can then detect this rule across familiar and novel speech items.

The children in the current study obtained overall lower accuracy scores compared to the adults in Peña et al. (2002) and Frost and Monaghan (2016), with both studies reporting an accuracy rate of around 70% in the moved-syllable condition with pauses (d = 1.307) and in the novel-syllable condition without pauses (d = 1.209) respectively.

Nonetheless, we obtained an effect size of 0.294 in the condition with pauses added to the familiarisation stream, which can be considered a small but meaningful effect of pauses. The smaller effect size may be attributed to the fact that adults are better explicit learners than children (Ferman & Karni, 2010; Ojima & Okanoya, 2020). In addition, the children watched a silent animation movie while listening passively to the familiarisation stream, while the adults in Peña et al. (2002) and Frost and Monaghan (2016) did not perform any other task during the listening phase. We believe that for our age group, engaging in active listening to the stream for 10 minutes is not feasible. However, this may have hindered the active learning process. It should also be noted that even though we found a significant increase in scores when pauses were added to the stream, our data showed that not all children in our study were able to learn the NAD rules. We found a significant positive correlation between children's age and the number of correct items, indicating that older children show more evidence of learning NADs. This age-related increase in performance may be due to the learning mechanisms that are used by the children in this task, reinforcing the idea that explicit learning performance increases with age. Older children may have more explicit metalinguistic knowledge (see Bialystok, 1986), which may have been beneficial in the present task. This supports previous findings by, for example, Ferman and Karni (2010) and Ojima and Okanoya (2020), who found more rule-learning success among adults and older children in artificial grammar learning paradigms when using a task that requires more explicit knowledge. It is also important to note that older children may simply have been better at understanding the task at hand. Even though there is no evidence to believe otherwise, we have not explicitly tested whether the children in our study understood the task, for example, by giving them the same task with familiar natural stimuli. It would be useful to do this in future studies, to be able to more reliably conclude that older children were better at learning NADs.

Conclusion

The results of this study strongly suggest that 7- to 11-year-old children have a better chance at learning artificial NADs when pauses are present in the speech stream. These pauses help divide the continuous stream into smaller chunks, making it easier to detect regularities within those chunks. The results also show that older children were more successful at detecting NADs than younger children, but that, overall, the children in our study were not as successful as the adults in previous studies based on the results of a forced-choice task. Nevertheless, performance was enhanced when prosodic cues, in the form of pauses, were added to the familiarisation stream. Only then were children able to discover the underlying NAD-rules across test items containing both moved and novel intervening elements. This reinforces the idea that prosodic cues facilitate language learners of all ages in discovering grammatical rules.

References

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear Mixed-Effects Models Using S4 Classes (R Package Version 0.999999-0).

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 57*(2), 498–510. https://doi.org/10.2307/1130604

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer. Version 6.0.37. Retrieved from http://www.praat.org/.

de Diego-Balaguer, R., Rodríguez-Fornells, A., & Bachoud-Lévi, A.-C. (2015). Prosodic cues enhance rule learning by changing speech segmentation mechanisms. *Frontiers in Psychology, 6,* 1478. https://doi.org/10.3389/fpsyg.2015.01478

Endress, A. D., & Bonatti, L. L. (2007). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition, 105*(2), 247–299. https://doi.org/10.1016/j.cognition.2006.09.010

Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language, 60*(3), 351–367. https://doi.org/10.1016/j.jml.2008.10.003

Ferman S., & Karni, A. (2010). No childhood advantage in the acquisition of skill in using an artificial language Rule. *PLoS One 5*(10): e13648. https://doi.org/10.1371/journal.pone.0013648

Frost, R. L. A., & Monaghan, P. (2016). Simultaneous segmentation and generalisation of non-adjacent dependencies from continuous speech. *Cognition, 147,* 70–74. https://doi.org/10.1016/j.cognition.2015.11.010

Grama, I. C., Kerkhoff, A., & Wijnen, F. (2016). Gleaning structure from sound: The role of prosodic contrast in learning non-adjacent Dependencies. *Journal of Psycholinguistic Research*, *45*(6), 1427–1449. https://doi.org/10.1007/s10936-016-9412-8

Grama I. C., & Wijnen, F. (2018). Learning and generalizing non-adjacent dependencies in 18-month-olds: A mechanism for language acquisition?. *PLoS* One *13*(10), e0204481. https://doi.org/10.1371/journal.pone.0204481

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package:

Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82*(1), 1–26. https://doi.org/10.18637/jss.v082.i13

Lammertink, I., Witteloostuijn, M. V., Boersma, P., Wijnen, F., & Rispens, J. (2019). Auditory statistical learning in children: Novel insights from an online measure. *Applied Psycholinguistics, 40*(2), 279–302. https://doi.org/10.1017/S0142716418000577

Long, J.A. (2020). jtools: Analysis and Presentation of Social Scientific Data. R package version 2.1. https://cran.r-project.org/package=jtools

Marchetto, E., & Bonatti, L. L. (2013). Words and possible words in early language acquisition. *Cognitive Psychology,* 67(3), 130-150. https://doi.org/10.1016/j.cogpsych.2013.08.001

Marchetto, E., & Bonatti, L. L. (2015). Finding words and word structure in artificial speech: the development of infants' sensitivity to morphosyntactic regularities. *Journal of Child Language, 42*(4), 873–902. https://doi.org/10.1017/S0305000914000452

Marimon, M., Hofmann, A., Veríssimo, J., Männel, C., Friederici, A. D., Höhle, B., & Wartenburger, I. (2021). Children's learning of non-adjacent dependencies using a web-based computer game setting. *Frontiers in Psychology,* 12, 734877. https://doi.org/10.3389/fpsyg.2021.734877

Morgan, J. L. (1986). *From Simple Input to Complex Grammar*. MIT Press. https://mitpress.mit.edu/books/simple-input-complex-grammar

Mueller, J. L., Bahlmann, J., & Friederici, A. D. (2008). The role of pause cues in language learning: The emergence of event-related potentials related to sequence processing. *Journal of Cognitive Neuroscience, 20*(5), 892–905. https://doi.org/10.1162/jocn.2008.20511

Mueller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences, 109*(39), 15953–15958. https://doi.org/10.1073/pnas.1204319109

Mueller, J. L., Milne, A., & Männel, C. (2018). Non-adjacent auditory sequence learning across development and primate species. *Current Opinion in Behavioral Sciences, 21,* 112–119. https://doi.org/10.1016/j.cobeha.2018.04.002

Ojima, S. & Okanoya, K. (2020). Children's learning of a semantics-free artificial grammar with center embedding. *Biolinguistics, 14,* 21-48.

Pacton, S., & Perruchet, P. (2008). An attention-based associative account of adjacent and nonadjacent dependency learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(1), 80–96. https://doi.org/10.1037/0278-7393.34.1.80

Paul, M., Männel, C., van der Kant, A., Mueller, J. L., Höhle, B., Wartenburger, I., & Friederici, A. D. (2021). Gradual development of non-adjacent dependency learning during early childhood. *Developmental Cognitive Neuroscience, 50,* 100975. https://doi.org/10.1016/j.dcn.2021.100975

Peña, M., Bonatti, L. L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*(5593), 604–607. https://doi.org/10.1126/science. 1072901
Perruchet, P., Tyler, M. D., Galland, N., & Peereman, R. (2004). Learning nonadjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology. General, 133*(4), 573–583. https://doi.org/10.1037/0096-3445.133.4.573

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory and Language, 35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Sánchez-Meca, J., Marín-Martínez, F., & Chacón-Moscoso, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological Methods, 8*(4), 448-467.

Schaadt, G., Paul, M., Muralikrishnan, R., Männel, C., & Friederici, A. D. (2020). Seven-year-olds recall non-adjacent dependencies after overnight retention. *Neurobiology of Learning and Memory, 171,* 107225. https://doi.org/10.1016/j.nlm.2020.107225

Skeide, M. A., & Friederici, A. D. (2016). The ontogeny of the cortical language network. *Nature Reviews Neuroscience, 17*(5), 323–332. https://doi.org/10.1038/nrn.2016.23

Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. W. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic infants. *Journal of Memory and Language, 49*(2), 249–267. https://doi.org/10.1016/S0749-596X(03)00024-X

Soderstrom, M., White, K.S., Conwell, E. and Morgan, J.L. (2007). Receptive grammatical knowledge of familiar content words and inflection in 16-month-olds. *Infancy*, *12*(1), 1-29. https://doi.org/10.1111/j.1532-7078.2007.tb00231.x

van der Kant, A., Männel, C., Paul, M., Friederici, A. D., Höhle, B., & Wartenburger, I. (2020). Linguistic and non-linguistic non-adjacent dependency learning in early development. *Developmental Cognitive Neuroscience, 45*, 100819. https://doi.org/10.1016/j.dcn.2020.100819

Zhang, Z., & Yuan, K. H. (2018). *Practical Statistical Power Analysis Using Webpower and R* (Eds). Granger: ISDSA Press.

Data, code and materials availability statement

Materials, data, and scripts are available online: https://osf.io/e43hw/

Ethics statement

The research was conducted in accord with the Research Ethics Code of the Leiden University Centre for Linguistics. The parents of all the children participating in the study signed an informed consent form that was sent to them through the schools that participated in the research project.

Authorship and Contributorship Statement

Van der Klis and Van Lieburg designed the experiment, collected the data, performed the statistical analyses, and drafted and revised the manuscript. Cheng and Levelt were responsible for the conception of the study, edited the manuscript and supervised the study. All authors are responsible for the interpretation of the results. All authors approved the final version of the manuscript and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acknowledgements

This work was supported by the Research Traineeship Programme of the Faculty of Humanities, Leiden University. We would like to thank the children for their participation, and Daltonschool De Margriet in Rotterdam and Haanstra primary school in Leiden for their cooperation. We thank Anouk Vinders for her help at the Daltonschool De Margriet. We also want to express our gratitude to Dr. Leticia Pablos Robles for offering us extensive statistical advice.

License