Difficulties in Nonadjacent Dependency Learning in French-Learning Toddlers

Julie Bodard, Thierry Nazzi, Géraldine Jean-Charles, Marco Pedrotti, Katrin Skoruppa

Difficulties in Nonadjacent Dependency Learning in French-Learning Toddlers.

Julie Bodard[a], Thierry Nazzi[b], Géraldine Jean-Charles[a], Marco Pedrotti[c] and Katrin Skoruppa[a]

[a] Institut des Sciences Logopédiques, Maison des Sciences du Langage et de la Communication – Université de Neuchâtel, Neuchâtel, Switzerland

[b] Integrative Neuroscience and Cognition Center, CNRS – Université Paris Cité, Paris, France

[c] Haute École Arc Santé, HES-SO University of Applied Sciences and Arts Western Switzerland – Delémont, Switzerland

Author Contributions

Conceptualization, J.B., T.N. and K.S.; Data curation, J.B. and G.J.-C.; Formal analysis, J.B., G.J.-C. and K.S.; Funding acquisition, J.B.; Investigation, J.B.; Methodology, J.B., T.N. and K.S. ; Project administration, J.B. and K.S.; Resources, J.B. and K.S.; Software, J.B., G.J.-C., M.P. and K.S.; Supervision, T.N. and K.S.; Visualization, J.B., G.J.-C. and K.S.; Writing – original draft, J.B., T.N. and K.S.; Writing – review and editing, J.B., T.N., G.J.-C., M.P. and K.S.

Correspondence

Correspondence to Julie Bodard: julie.bodard@unine.ch

Disclosure

The authors declare that no competing financial or non-financial interests existed at the time of publication.

Abstract

Purpose: In the past three decades, statistical learning, that is, the capacity to detect patterns and regularities in the environment, has been shown to have an important role in language development. In particular, the ability to detect nonadjacent dependencies (NADs) between linguistic elements that are separated by intervening material seems to be linked to morphosyntactic development. Yet, there are few studies with French-speaking children, and none using artificial language learning methodologies.

Method: We investigated the acquisition of nonadjacent dependencies in 18-month-old (n = 30) and 27-month-old (n = 32) monolingual French-learning children in an adaptation of Gómez (2002) artificial grammar learning study to a central fixation eye-tracking paradigm. After a 2-minute familiarization with three element strings (e.g. /na sokɛ fib/), we monitored looking times for grammatical versus ungrammatical strings in two blocks of 4 test trials each, expecting a novelty effect at 18 months and exploring the behavior at 27 months.

Results: No significant effect of grammaticality was found at either 18 or 27 months for the overall task. However, at 27 months we observed a significant decrease in gaze duration for ungrammatical test trials between the first and the second block, leading to the emergence of a marginally significant preference for grammatical stimuli in the second block, that could potentially indicate the start of novel NAD learning.

Conclusions: Given the failure to clearly replicate effects observed in the studies with English-learning children, we propose several theoretical hypotheses to account for our results and discuss the importance of differences in age and language background as well as methodological parameters— a well-known challenge in research with young children. Further research is needed to develop robust tasks and to better understand the developmental trajectory of nonadjacent dependency processing abilities.

Keywords: Statistical Learning; Nonadjacent Dependency; Toddlers.

Introduction

Over the last three decades, research has underlined the importance of statistical learning, that is, the ability to detect patterns and regularities in the environment, for language acquisition (Saffran, 2020). In particular, it has been suggested that the ability to learn nonadjacent dependencies (NADs), i.e. regularities between co-occurring elements in a word, phrase or sentence separated by intervening material, might be important for morphosyntactic development (Erickson & Thiessen, 2015). NADs are observed in many languages, for example in number subject-verb agreement in English (e.g., ‘the girl sleeps’ vs. ‘the girls sleep’) and in French (“la fille dort” [la fij dɔʁ] vs. “les filles dorment” [lɛ fij dɔʁm]). In a seminal study, Gómez (2002) found that 18-month-old English-learners rapidly learn the statistical regularities of nonadjacent elements in an artificial language, when the variability of the item that separates them is large enough during familiarization (X = 24 *vs* 3 and 12). In a head-turn preference procedure (HPP), participants were exposed to one of two artificial languages for three minutes. Strings were made up of sets of three nonwords with pauses in between (e.g., pel-kicey-jic), consisting of two nonadjacent dependencies between the first and last items (L1: aXd, bXe; L2: aXe, bXd). Then, during test, they heard strings following or violating these dependency rules. A novelty effect was observed, with longer listening times for the strings violating the regularity, establishing novel NAD learning in infants for the first time. The present study was conducted to contribute new data regarding the generalizability of this learning mechanism, testing infants at different ages and learning a different language, French, on which no data exists for novel NAD learning.

Further studies on NAD learning in English revealed developmental changes. For example, Gómez & Maye (2005) extended Gómez (2002) by testing toddlers of different ages with NADs in artificial grammars. They observed no preference at 12 months, a familiarity preference at 15 months and a novelty preference at 17 months. Based on Hunter & Ames (1988) model, they interpreted these results as reflecting changes in the toddler’s ability to track NADs, capturing its emergence between 12 and 17 months, and its increase in efficiency between 15 and 17 months.

A few studies also tested whether NAD learning extends to other languages. Using similar stimuli and procedures as Gómez (2002), NAD learning attested by a novelty preference was found in Dutch-learning 18-month-olds (Kerkhoff et al., 2013). Further studies on German, measuring event-related potentials (ERPs), suggest that toddlers can learn novel NADs from as early as 3 months up to 2 years, while this ability appears to decline by around 2 to 3 years (Mueller et al., 2019; van der Kant et al., 2020). It has been proposed that this developmental shift might be driven by increasing knowledge of the NADs in the native language, which may interfere with novel NAD learning (Paul et al., 2021). These studies establish NAD learning beyond English, although further additional crosslinguistic variation from more diverse, non-Germanic languages is needed.

The need for further crosslinguistic extensions is further motivated by studies on native language NAD comprehension, which have shown crosslinguistic differences in developmental trajectories. In particular, studies on the acquisition of subject-verb agreement in French have established early comprehension, by 30 months of age, of at least one agreement subsystem (based on liaison consonant, e.g., plural : "**le**/**les** garçon(s) i**l**/il**s-**embrassent" [**lə**/**le**...i**l**/il**z-**ɛ̃bras]/'the boys they kiss'; Legendre et al., 2010), while at a comparable age and using the same methods, English learners showed no evidence of comprehension, and Spanish learners only partial success (Legendre et al., 2014). The authors suggest that these cross-linguistic differences in the age of comprehension of agreement systems are linked to their perceptual saliency and cue reliability, and further studies suggest this early comprehension in French-learning toddlers might be linked to early development of a sensitivity to morphosyntactic nonadjacent regularities in their native language. Indeed, French-learning toddlers become sensitive to NADs involving subject-verb agreement during the second year of life, tracking remote dependencies between determiners and auxiliaries by 17 months (e.g., "**la/les** coupile(s) **va/vont** bientôt conduire" [**la**/**le** ... **va**/**vɔ̃**...]/'The$_{SG/PL}$ X will$_{SG/PL}$ soon drive'; van Heugten & Shi, 2010) and between determiners and number marking on verbs at 18 and 24 months (e.g., "**le**/**les** garçons lit/li**sent**" [**lə**/**le**...li/li**z**]/'the boy(s) read(s)'; Nazzi et al., 2011). They also show sensitivity to the dependency between the determiner in the noun phrase and the redundant subject pronoun, functioning as verbal affix agreement (e.g., singular : “**le** garçon i**l-**arrive”/’the boy he walks’, pronounced [lə...i**l**ariv] with the liaison [il-a] ; plural : “**les** garçons il**s-**arrivent”/’the boys they walk’, pronounced [le...il**z**ariv] with the liaison [ilz-a]) at six different ages between 14 and 24 months (Culbertson et al., 2016). Importantly, Culbertson et al. (2016) reported several alternations of preferences across age for grammatical and ungrammatical stimuli, with a novelty effect in 15-, 18- and 24-month-olds, but a familiarity effect in 14- and 21-month-olds (and no effect in 20-month-olds). These changes of patterns were interpreted as evidence for distinct stages in the acquisition of these nonadjacent dependencies (from reflecting surface statistics to higher-level morphological features), revealing a complex developmental trajectory.

Given these early native morphosyntactic acquisitions in French, novel NAD learning in French toddlers might differ from what has been observed in English, Dutch and German, a gap that the present study was designed to address. Following Gómez (2002), we investigated whether French-learning toddlers aged 18 and 27 months can learn novel NADs of non-identical elements (AXB) in artificial grammars, using a central fixation paradigm with a gaze-contingent display. If this ability found in English- and Dutch-learning 18-month-olds generalizes, we expect French-learning 18-month-olds to learn these novel NADs, which would be attested by a group-level novelty effect. However, given the lack of prior behavioral data at older ages, it is more challenging to predict whether 27-month-olds would show evidence of learning, and if so, the direction of this effect.

Method

Participants

Participants were recruited in the Canton of Neuchâtel in Switzerland. Families were living in French-speaking Switzerland, and French was the only language spoken at home. According to parental reports, all children were born at term (>37 weeks) and had no known sensory, anatomical, cognitive or neurological conditions.

At 18 months, data from twenty-nine monolingual French-learning children (15 females, 14 males) aged 18;0 to 19;20 months (M = 18;18 months) were included in the analyses. An additional nine children participated but were excluded due to fussiness (n = 3), lack of interest (n = 1) or technical problems (n = 4), or extremely short looking times (n = 1, outlier with a mean looking time inferior to -1.5 interquartile range).

At 27 months, data from thirty-one monolingual French-learning children (15 females, 16 males) aged 26;18 to 28;2 months (M = 27;2 months) were included in the analyses. Twenty-three of the children contributed data at both ages. Eight children contributed data only at 27 months; of these, six had participated at 18 months but their data were excluded, and two new participants were included at 27 months. An additional five children participated but were excluded due to fussiness (n = 1), refusal to perform the task (n = 1), technical problems (n = 2), or extremely short looking times (n = 1, same outlier, under the same conditions as at 18 months).

As part of a larger longitudinal project[1], this study was approved by the local ethics committee on research involving humans (Swiss Ethics section of the canton of Vaud, number 2018-00765). The participants were recruited through contacts with nurseries and pediatricians, by advertisement on social networks and press and word of mouth. Their parents gave their free and informed consent to participate in this study via a detailed form.

Stimuli

Auditory stimuli. The auditory strings, adapted to French from Gómez (2002), were composed of three-element strings with NADs, following the structures [aXc] and [bXd]. The first and third elements were fixed monosyllables, whereas the middle element varied among 24 disyllables. We built 4 artificial grammars for the familiarization phase with 2 lexicons, and for each lexicon 2

[1] This study is part of a longitudinal follow-up exploring links between statistical learning and morphosyntactic development in 18-to-40-month-old children. These data come from the first two of three phases. Participants were exposed to this novel NAD learning task in artificial language at 18 and 27 months of age, in the same place and with the same setting. In both sessions, an additional task was proposed along this NAD sensitivity task in artificial language: an adjacent-dependency sensitivity task in artificial language at 18 months, and a NAD sensitivity task in native language at 27 months. Each task, including the one presented here, lasted around 5 min, depending on the child's interest. A short break was made between the two tasks when needed, taken outside of the sound-attenuated booth, depending on the child's behavior and the parents' input. The procedure of the three tasks were rather similar, however the native language task had no familiarization phase and different stimuli. The order in which the tasks were performed was counterbalanced according to two randomizations.

different dependency arrangements (see Table 1). Thus, participants were familiarized with one grammar (48 strings: 2 NADs * 24 middle elements). Then, during the test phase, they were presented with previous strings respecting the familiarized patterns (grammatical) and new strings from the other, unfamiliar grammar (ungrammatical). Note that Grammar 1 grammatical stimuli were Grammar 2 ungrammatical stimuli and inversely, and Grammar 3 grammatical stimuli were Grammar 4 ungrammatical stimuli and inversely.

Place Table 1 here

The same 6 grammatical and 6 ungrammatical strings were combined in different orders to form 4 grammatical and 4 ungrammatical test trials, for a total of 8 test trials for each group (see full list of strings in Appendix A).

For the familiarization and test stimuli, a female speaker native from the canton recorded several tokens of each nonword (e.g. /na/, /sokɛ/, /fib/) in a sound attenuated booth, with a studio microphone (Rode NT1-A with pop shield) and a Steinberg audio interface (UR22 mkii), using Audacity version 2.2.2 (Audacity Team, 2018). Individual word tokens were assembled together to constitute the 'strings' (e.g. /na sokɛ fib/), with 200 ms pauses between nonwords and 850 ms pauses between 'strings'. String durations were 1 951 ms in average (min = 1 591 ms, max = 2 592 ms), for a total trial length of 16 s.

Visual stimuli and motivational sounds. A child-friendly Winnie the Pooh wallpaper and music were chosen to attract children into the booth. Two muted child-friendly cartoons (involving the character Pocoyo) were chosen for the familiarization phase, during which they were displayed on the whole screen. An animated bee or butterfly gif (height: 282 px, width: 177 px) was selected as an attention grabber for the beginning of each trial, to be played together with 2 nursery rhyme melodies played by music boxes. For the test phase, a geometrical colored animation (heigth: 490 px, width: 480 px) was chosen to be played at the center of the screen, in parallel with the auditory strings.

In order to avoid exposing our participants to the same stimuli and orders twice, four pseudo-random groups were made for the first session of the study (distribution by order of entry in the study and homogenization of the number of participants per group) and kept for the second session. Grammar, order of the items, repartition of the test trials regarding their grammaticality and different visual stimuli were counterbalanced between the 4 groups within the task and between the two age sessions, to control for potential biases.

Procedure and Apparatus

The children came to the Babylab of the University of Neuchâtel with one or both parents, or a familiar caregiver. They were installed in a sound-attenuated booth on their caregiver's laps at 60 cm from a screen (24'', 1920 x 1080 pixels) placed on a small table, on which a Tobii Pro X3-120 eye tracker with External Processing Unit was mounted. Auditory stimuli were played through hidden loudspeakers (Fostex PM0.5n, 70W amplifier) under the screen. Caregivers heard masking multi-speakers babble through headphones. They were also asked to not look at the screen and to keep their heads tilted forward or their eyes closed, so that they were blind to the visual stimuli and the eye tracker could not catch their gaze. A camera at the top of the screen allowed the experimenter to monitor the progress of the test online and to record the behaviors of the children and parents for offline checking.

We used a central fixation paradigm with gaze-contingent display, using eye gaze as a proxy measure for auditory attention. For most test sessions[2], the task was run using Matlab version 2019b (MathWorks, 2019) using the Psychtoolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) and the Titta toolbox (Niehorster et al., 2020).

Children were first exposed to a 5-point calibration with pulsating dots, manually validated. An overview of the study procedure that followed can be found in Figure 1. During the initial familiarization phase, children heard 48 strings, while watching a silent cartoon. This phase lasted 2

[2] We initially ran the task with E-Prime (Version 2.0 ; Schneider et al., 2010), which was used to test the first 10 18-month-olds (and to calibrate all 18-month-olds), but we then decided to switch software due to technical issues, maintaining all task parameters identical.

min. 12 s. During the test phase, 8 test trials (4 grammatical, 4 ungrammatical) were played along with the moving geometrical form, with a maximal duration of 2 min. 8 s. Before each test trial, a gaze-contingent attention grabber was played at the center of the screen until the children looked at it for more than 1 second, which would trigger the test trial. Test trials either lasted for 16 seconds or ended when the child stopped looking for 2 consecutive seconds. At the end of the experiment, an animated clapping emoji and a clapping sound were presented.

Place Figure 1 here.

Data Processing and Analysis

We developed a Matlab script (MathWorks, 2019) to extract and process the data. The raw data provided the position of the gaze on the screen every 8 ms, thus we obtained gaze "hit" or "no hit" in the area of interest (AOI) i.e. the visual stimulus (AOI of test trials visual stimulus: X: left=576 px, right=1344 px, Y: top=324 px, bottom=756 px). We considered short tracking losses (no hit < 88 ms) and blinks (no hit between 88 ms and 304 ms) as continuous gaze, whereas longer gaze absences (no hit > 304 ms) were considered as absence of gaze. We manually checked the videos and excluded trials during which problems occurred and performed an interrater reliability check (see details below). A second judge checked 15% of the videos for exclusion or inclusion. Substantial agreement (according to Landis & Koch (1977) in Cohen's kappa tests was found at 18 months (95% of agreement, k = 0.68) and almost perfect agreement at 27 months (97.5% of agreement, k = 0.89).

Results

At 18 months, twenty-nine test trials (of 232, thus 12,5%) were removed, because of (a) extremely short looking times (< 2 sec, n=12) (b) fussiness or crying (n=6), (c) eye-tracking problems (n=9), (d) caregiver's intervention (n=1), (e) technical problem (n=1). Thus, 203 valid test trials remained for analysis, and each participant contributed at least two valid test trials per condition. At 27 months, thirty-seven test trials (of 248, thus 14,9%) were removed, because of (a) extremely short looking times (< 2 sec, n=21) (b) fussiness or crying (n=3), (c) eye-tracking problems (n=6), (d) caregiver's intervention (n=4), (e) interrupted trial/no data (n=3). Thus, 211 valid test trials remained for analysis, and each participant contributed at least two valid test trials per condition (all raw data can be found in Supplemental material S1). Results are presented in Figure 2, left panel, for the data at 18 months, and in Figure 2, right panel, for the data at 27 months. Note that at both ages, participants had long looking times, especially in the first block (they looked during the entire trial in 33,6% and 42% of trials at 18 and 27 months respectively).

Place Figure 2 here.

Since raw looking times were not normally distributed (Shapiro-Wilk test: $W = 0.83$, $p < .001$), we transformed them into proportion scores by dividing each child's looking time during a trial by the maximum possible duration of a trial, and build generalized linear mixed models for beta distributions, with a logit link, using the glmmTMB package (Brooks et al., 2017; version 1.1.10) in R 4.3.3 (R Core Team, 2024), to analyze the effect of stimulus grammaticality on looking time per trial and per participant. The model included gaze proportion as dependant variable, grammaticality (grammatical vs. ungrammatical trials) and block (block 1 vs. 2) as fixed effects, as well as the interaction of the latter, moderated by age as nested factor (18 vs. 27 months) and the variable participant as random effect (R formula : glmmTMB(gaze_prop ~ age / (condition*block) + (1 | subject), data = data, family = beta_family(link = "logit")).

Place Table 2 here

All results of the GlmmTMB model are shown in Table 2. At 18 months, we found a significant effect of block at 18 months ($\beta = -.799$, $p < .001$), indicating that the 18-month-olds looked significantly less in Block 2 than in Block 1 (Block1: M = 12.15 s vs. Block2: M = 8.89 s). At 27 months, we found a significant interaction between condition and block at 27 months ($\beta = -.671$, $p = .034$). To follow up on this interaction, we performed post-hoc Dunn's tests targeting 4 comparisons between blocks and conditions. They showed no significant difference in looking times between grammatical and ungrammatical trials in Block 1 (Block1-Gram: M = 12.06 s., Block1-Ungram: M = 13.22 s., $z = -1.02$, p

= .153 uncorrected, p = .204 with Benjamini-Hochberg correction), yet a marginal preference for grammatical over ungrammatical trials in block 2 (Block2-Gram: M = 11.96 s., Block2-Ungram: M = 10.52 s., z = 1.73, p = .042 uncorrected, p = .084 with Benjamini-Hochberg correction). They also showed a significant decrease in looking times for ungrammatical trials between block 1 and block 2 (z = 2.99, p = .001 uncorrected, p = .006 with Benjamini-Hochberg correction), whereas there was no significant difference for grammatical trials (z = 0.02, p = .494 uncorrected, p = .494 with Benjamini-Hochberg correction).[3]

Discussion

Investigating the learning of novel nonadjacent dependencies in monolingual French-learning toddlers in artificial grammar, we expected a novelty effect indicating easy acquisition in 18-month-olds, based on Gómez (2002) and Kerkhoff et al. (2013). We also aimed to explore the performance of 27-month-olds, for whom literature is scarcer. However, we found no significant grammaticality effect in either age group. At 18 months, we observed an expected significant drop in interest between the first and second halves of the task, which was not modulated by grammaticality. Interestingly, however, at 27 months, interest declined between the first and second halves of the task for ungrammatical trials, while interest in grammatical trials remained stable. This differential decline led to a marginally significant grammaticality effect in the second block. We yet failed to establish robust novel NAD learning in French-learning 18- and 27-month-olds, even though several studies have established that French-learner toddlers are sensitive to nonadjacent dependencies of fairly similar structures in their native language by 14 to 24 months (van Heugten & Shi, 2010; Nazzi et al., 2011; Culbertson et al., 2016).

A first possible interpretation of our results is that at both ages, our French-learning participants did not succeed in learning novel non-adjacent dependencies because of methodological

[3] When data of the outlier participant are included, the results profile is similar with weaker effects in Dunn's tests between grammatical and ungrammatical trials in Block 2 (p = .058 uncorrected, p = .115 with Benjamini-Hochberg correction) but still a significant decrease between block 1 and block 2 for ungrammatical trials (p = .002 uncorrected, p = .009 with Benjamini-Hochberg correction).

reasons. Indeed, our task differs from Gomez's in terms of procedure (central fixation *vs.* HPP) and visual stimuli (graphic animation *vs.* lights), which may have hindered NAD learning. Our choice of a visual fixation procedure and visual stimuli may have increased participants' interest in the task, leading to a relatively low drop-out rate (18 months: 12%, 27 months: 11%), but also may contribute to our globally long looking times. As noted by the ManyBabies Consortium (2020) and Junge et al. (2020), HPP generally leads to higher drop-out rates, but also larger familiarity effects, potentially due to the need for a motor response (headturn) increasing sensitivity. Furthermore, we chose stimuli that are visually attractive to toddlers: a silent cartoon for familiarization, an animal animation as attention grabber and colorful moving forms as visual stimulus during the test phase. Colors and motion are known to attract visual attention (Atkinson & Braddick, 2013), and animated characters, particularly anthropomorphic animals, have proven effective for educational activities (Snyder & Mares, 2021; Valkenburg & Vroone, 2004). Consequently, our visually engaging stimuli may have heightened the toddlers' visual interest compared to HPP procedures using blinking lights. This could explain why our participants showed overall high interest in test trials, especially in the first block.

Our visually attractive stimuli may also have interfered with the processing of auditory stimuli, potentially leading to a lack of NAD learning. The use of cartoons during the familiarization phase may have decreased participants' attention to the auditory stimuli, however, according to Mueller et al. (2019) this is unlikely due to the prevalence of auditory over visual attention (auditory bias) in toddlers. Animal attention grabbers may also have triggered implicit or explicit naming effects (Helo et al., 2017). In any case, our stimuli were not only more salient but also more complex, which may have increased the cognitive demands of the task. According to Hunter & Ames (1988)'s model, when tasks are more complex, a longer familiarization phase (and we would also add longer test trials to avoid ceiling effects) may be required to allow complete learning and allow for discrimination NADs. This hypothesis may align with the marginally significant preference for grammatical stimuli in the second block in 27-month-olds.

A second, more developmental interpretation of our results is that the lack of preference shown by our French-learning toddlers is related to the ages that we tested and may represent normal phases in novel NAD learning. This idea relies on the model of cyclical learning development proposed by Culbertson et al. (2016), who found that, in a task assessing native NAD knowledge, toddlers of different age groups display varying preferences, interspersed with phases of no preference (Hunter & Ames, 1988). A familiarity preference would reflect incomplete information processing, whereas a novelty preference would signal robust knowledge, with an absence of preference being found during the passage from one to the other. Culbertson et al. (2016) further described two cycles of familiarity/novelty preference, reflecting two levels of acquisition: an initial level focused on surface sound patterns and a subsequent level at a more abstract morphological feature level. Thus, the absence of preference we observed may indicate that our participants are between two preference phases and possibly in two different acquisition levels (surface vs. feature). If so, novel NAD learning by French-learning infants would happen at a different pace than in English (Gómez, 2002; Gómez & Maye, 2005), in line with studies having shown that children's and adult's entrenched expectations about speech co-occurrences from their native language influence what they learn from novel linguistic input (Frost et al., 2019). Our inconclusive results at 27 months might also be linked to the decline in toddlers' ability to learn NADs in an artificial language under passive listening conditions in ERP studies between 2 and 3 years of age (Mueller et al., 2019; Paul et al., 2021; van der Kant et al., 2020).

A third possible interpretation of our findings is that the marginal effect of grammaticality at 27 months in the second block of the experiment might be a weak sign that these older children have started to learn the NADs in the task. Indeed, we observed a significant drop in interest for unfamiliar stimuli between the two blocks of the experiment, which may suggest a form of discrimination between the stimuli, and thus might attest that toddlers at this age were in the process of learning the novel NADs. This effect suggests that infants were developing a familiarity preference in this second block (which however was only marginal), highlighting the difficulty of the task (following

Hunter & Ames, 1988; Culbertson et al., 2016). The reason these effects appear along the test phase might be due to infants requiring more time to learn than the familiarization time given to them, and/or due to long listening times during the first block.

In conclusion, we did not find clear evidence of novel NAD learning in French-learning toddlers at either 18 or 27 months, failing to extend such learning abilities previously demonstrated in English, Dutch and German. Our results at 27 months could be taken as an indication of ongoing learning, but the evidence remains inconclusive at this point. Future work will have to reevaluate novel NAD learning in French, but also in different languages, extending the age range tested to get a fuller picture of the developmental trajectory of this learning ability across languages and specify the linguistic parameters that might modulate this trajectory. In doing so, parallel cross-linguistic studies using similar methods and stimuli could be conducted, adjusting the methodology used in the present study based on the considerations discussed above, and increasing sample size given that, as in all developmental studies, collapsing data across children may be affected by possible individual differences (Roder et al., 2000). Importantly, it is now acknowledged that early speech and language processing abilities partly predict later language outcomes (see for example, Cristia et al., 2014, for vocabulary acquisition), and it has been proposed that the ability to learn nonadjacent dependencies might predict morphosyntactic acquisition such as the agreement of verbal and nominal groups (Erickson & Thiessen, 2015). If this link were confirmed, then we could develop tools to measure children's early ability to learn statistical regularities, and identify infants and toddlers at risk of language delay, in a context in which the need for early intervention is widely admitted (Fricke et al., 2013; Guralnick, 2011), but the means to identify at risk children are still lacking (e.g., Bishop et al., 2017). Setting up methodologically-sound and robust tasks to assess morphosyntactic skills, such as the learning of NADs, at an individual level, would be a great advancement for research and for the speech and language therapy clinic.

Acknowledgments

This project was supported by an SNF Doc.CH subside (FN P0NEP1_178712) and the Chaire de logopédie I at the Institut des sciences logopédiques, Université de Neuchâtel. The authors would like to thank all participants and their families for taking part in this research, as well as daycare staff and pediatricians, who helped with recruitment. Special thanks to Salomé Schwob who registered the auditory stimuli, Dahliane Labertonière, Salomé Schwob and Letizia Volpin who helped out for testing in a difficult time, and Andri Manser from LiRI statistical consultation of the University of Zürich.

Data Availability Statement

The data presented in this study are included in the supplemental material (S1) with a “Read me” tab (S2), further inquiries can be directed to the corresponding author.

References

Atkinson, J., & Braddick, O. (2013). Visual Development. Dans P. D. Zelazo (Éd.), *The Oxford Handbook of Developmental Psychology, Vol. 1 : Body and Mind* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199958450.013.0010

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & Consortium, and the C.-2. (2017). Phase 2 of CATALISE : A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry*, *58*(10), 1068‑ 1080. https://doi.org/10.1111/jcpp.12721

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433‑ 436. https://doi.org/10.1163/156856897X00357

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Machler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378- 400. https://doi.org/10.3929/ethz-b-000240890

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting Individual Variation in Language From Infant Speech Perception Measures. *Child Development*, *85*(4), 1330- 1345. https://doi.org/10.1111/cdev.12193

Culbertson, J., Koulaguina, E., Gonzalez-Gomez, N., Legendre, G., & Nazzi, T. (2016). Developing knowledge of nonadjacent dependencies. *Developmental Psychology*, *52*(12), 2174- 2183. https://doi.org/10.1037/dev0000246

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language : Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66- 108. https://doi.org/10.1016/j.dr.2015.05.002

Fricke, S., Bowyer-Crane, C., Haley, A. J., Hulme, C., & Snowling, M. J. (2013). Efficacy of language intervention in the early years. *Journal of Child Psychology and Psychiatry*, *54*(3), 280- 290. https://doi.org/10.1111/jcpp.12010

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research : A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128- 1153. https://doi.org/10.1037/bul0000210

Gómez, R. L. (2002). Variability and Detection of Invariant Structure. *Psychological Science*, *13*(5), 431- 436. https://doi.org/10.1111/1467-9280.00476

Gómez, R., & Maye, J. (2005). The Developmental Trajectory of Nonadjacent Dependency Learning. *Infancy*, *7*(2), 183- 206. https://onlinelibrary.wiley.com/doi/abs/10.1207/s15327078in0702_4

Guralnick, M. J. (2011). Why Early Intervention Works. *Infants and young children*, *24*(1), 6- 28. https://doi.org/10.1097/IYC.0b013e3182002cfe

Helo, A., van Ommen, S., Pannasch, S., Danteny-Dordoigne, L., & Rämä, P. (2017). Influence of semantic consistency and perceptual features on visual attention during scene viewing in toddlers. *Infant Behavior and Development*, *49*, 248- 266. https://doi.org/10.1016/j.infbeh.2017.09.008

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. *Advances in Infancy Research*, *5*, 69- 95.

Junge, C., Everaert, E., Porto, L., Fikkert, P., de Klerk, M., Keij, B., & Benders, T. (2020). Contrasting behavioral looking procedures : A case study on infant speech segmentation. *Infant Behavior and Development*, *60*, 101448. https://doi.org/10.1016/j.infbeh.2020.101448

Kerkhoff, A., Bree, E. D., Klerk, M. D., & Wijnen, F. (2013). Non-adjacent dependency learning in infants at familial risk of dyslexia. *Journal of Child Language*, *40*(1), 11- 28. https://doi.org/10.1017/S0305000912000098

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, *36*(14), 1- 16.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159- 174. https://doi.org/10.2307/2529310

Legendre, G., Barrière, I., Goyet, L., & Nazzi, T. (2010). Comprehension of Infrequent Subject–Verb Agreement Forms : Evidence From French-Learning Children. *Child Development*, *81*(6), 1859- 1875. https://doi.org/10.1111/j.1467-8624.2010.01515.x

Legendre, G., Culbertson, J., Zaroukian, E., Hsin, L., Barrière, I., & Nazzi, T. (2014). Is children's comprehension of subject–verb agreement universally late? Comparative evidence from French, English, and Spanish. *Lingua*, *144*, 21- 39. https://doi.org/10.1016/j.lingua.2013.05.004

ManyBabies Consortium. (2020). Quantifying Sources of Variability in Infancy Research Using the Infant-Directed-Speech Preference. *Advances in Methods and Practices in Psychological Science*, *3*(1), 24- 52. https://doi.org/10.1177/2515245919900809

MathWorks. (2019). *MATLAB* (Version Version R2019b) [Software]. The MathWorks, Inc. https://www.mathworks.com

Mueller, J. L., Friederici, A. D., & Männel, C. (2019). Developmental changes in automatic rule-learning mechanisms across early childhood. *Developmental Science*, *22*(1), e12700. https://doi.org/10.1111/desc.12700

Nazzi, T., Barrière, I., Goyet, L., Kresh, S., & Legendre, G. (2011). Tracking irregular morphophonological dependencies in natural language : Evidence from the acquisition of subject-verb agreement in French. *Cognition*, *120*(1), 119‑ 135. https://doi.org/10.1016/j.cognition.2011.03.004

Niehorster, D. C., Andersson, R., & Nyström, M. (2020). Titta : A toolbox for creating PsychToolbox and Psychopy experiments with Tobii eye trackers. *Behavior Research Methods*, *52*(5), 1970‑ 1979. https://doi.org/10.3758/s13428-020-01358-8

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics : Transforming numbers into movies. *Spatial Vision*, *10*(4), 437‑ 442. https://doi.org/10.1163/156856897x00366

R Core Team. (2024). *R: A language and environment for statistical computing* (Version 4.3.3) [Software]. R Foundation for Statistical Computing. https://www.R-project.org/

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. https://doi.org/10.1207/S15327078IN0104_9

Saffran, J. R. (2020). Statistical Language Learning in Infancy. *Child Development Perspectives*, *14*(1), 49‑ 54. https://doi.org/10.1111/cdep.12355

Schneider, W., Eschman, A., & Zuccolotto, A. (2010). E-prime, Version 2.0. 8.90 [computer software]. *Pittsburgh, PA: Psychology Software Tools*.

Snyder, M. N., & Mares, M.-L. (2021). Preschoolers' choices of television characters as sources of information : Effects of character type, format, and topic domain. *Journal of Experimental Child Psychology*, *203*, 105034. https://doi.org/10.1016/j.jecp.2020.105034

Valkenburg, P. M., & Vroone, M. (2004). Developmental Changes in Infants' and Toddlers' Attention to Television Entertainment. *Communication Research*, *31*(3), 288‑ 311. https://doi.org/10.1177/0093650204263435

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

van Heugten, M., & Shi, R. (2010). Infants' sensitivity to non-adjacent dependencies across phonological phrase boundaries. *The Journal of the Acoustical Society of America*, *128*(5), EL223‑ EL228. https://doi.org/10.1121/1.3486197

Tables and Figures

Table 1. *Grammar Structures and Test Trials Strings.*

Figure 1. *Overview of visual stimuli and task proceedings.* TT = test trial. AG = attention grabber.

Figure 2. *Distribution and median of looking times by grammaticality and block at 18 months (left panel) and 27 months (right panel).* The wider the violin plot, the more observations there are (small dots represent participant means). The tips of the whiskers in the box plots represent 1.5 interquartile range. Big red dots represent the median of the sample.

Table 2. *GlmmTMB results. CI = confidence interval. p = p-value.*

Appendix

A – List of audio stimuli

Supplemental Material

S1. Raw Data Table. *We provide a table with the raw looking times from the study per toddler and per test trials. The .xlsx file is named "S1_Looking_time_data"*

S2. Raw Data Table Description. *We provide a table with a description of the variables provided in the table. The .xlsx file is named* **"***S2***_***Read_me*".