BRIEF RESEARCH REPORT

Learning non-adjacent regularities at age 0;7[1]

JUDIT GERVAIN

CNRS and Université Paris Descartes, Sorbonne Paris Cité, Paris, France

AND

JANET F. WERKER

University of British Columbia, Vancouver, Canada

ABSTRACT

One important mechanism suggested to underlie the acquisition
of grammar is rule learning. Indeed, infants aged 0;7 are able to learn
rules based on simple identity relations (adjacent repetitions, ABB:
"wo fe fe" and non-adjacent repetitions, ABA: "wo fe wo", respectively;
Marcus *et al.*, 1999). One unexplored issue is whether young infants
are able to process both adjacent and non-adjacent repetitions. As
the previous studies always compared the two types of repetition
structures directly, the ability to learn only one of them was sufficient for
successful discrimination in these tasks. The present study reports
two experiments, in which we test the ability of infants aged 0;7 to
discriminate adjacent and non-adjacent repetition structures against
random controls (ABB vs. ABC and ABA vs. ABC). We show that,
contrary to some previous proposals, infants aged 0;7 successfully
discriminate both repetition types from random controls, but show no
spontaneous preference for either of them.

INTRODUCTION

Understanding the mechanisms underlying infants' ability to acquire
the grammar and vocabulary of their native language(s) has been central
in cognitive developmental psychology. Several mechanisms have been

---

860

proposed for different aspects of language acquisition. One mechanism, suggested to underlie the acquisition of grammar, is rule learning (Gómez & Gerken, 1999; Guasti, 2002; Marcus, Vijayan, Rao & Vishton, 1999). A particularly early demonstration of this ability was given by Marcus *et al.* (1999), who showed that infants aged 0;7 are able to learn rules based on identity relations (i.e. they discriminate adjacent repetitions in final from initial position, i.e. ABB: ''wo fe fe'' from AAB ''wo wo fe'', and more importantly for the current study, adjacent repetitions, i.e. ABB: ''wo fe fe'' from non-adjacent repetitions, i.e. ABA: ''wo fe wo''). This study initiated a large body of work exploring the nature of this learning mechanism (Endress, Nespor & Mehler, 2009; Frank, Slemmer, Marcus & Johnson, 2009; Gervain, Macagno, Cogoi, Pena & Mehler, 2008; Johnson *et al.*, 2009; Marcus, Fernandes & Johnson, 2007). However, several questions remain open.

One unexplored issue is whether adjacent and non-adjacent repetitions are processed similarly. The grammars of natural languages make use of both types of regularities (e.g. number agreement: The **boy is** *playing chess* vs. *The* **boy**, *whom I have never seen before*, **is** *playing chess*). It is, therefore, important to understand whether infants are able to learn both structures early on. Previous work (Gervain *et al.*, 2008; Gervain *et al.*, 2011) has shown that newborns can discriminate patterns based on adjacent repetitions (ABB) from random controls (ABC), but not non-adjacent repetitions (ABA) from random controls (ABC). Non-adjacent dependencies of co-occurring, but non-identical (i.e. not repetition-based) word classes (e.g. **aXc**: **pel** wadim **jic**, **pel** loga **jic**, etc.) have been found too difficult for infants aged 1;0 to learn (Gómez & Maye, 2005), unless given prior experience first with an adjacent dependency between the same word classes (Lany & Gómez, 2008). Successful performance on non-adjacent dependency learning between non-identical word classes starts at age 1;3 (Gómez & Maye, 2005). With natural language stimuli, sensitivity to non-adjacent dependencies seems to appear even later. English-learning infants aged 1;6, but not those aged 1;3, recognize the grammatical non-adjacent dependency between the auxiliary *is* and the verb ending *-ing*, and discriminate it from the ungrammatical *can …-ing* construct (Santelmann & Jusczyk, 1998). In Dutch, infants aged 2;0, but not those aged 1;5, have been found to track the non-adjacent dependency between the definite article *het* and the diminutive *-je*, highly frequent in corpora of infant-directed speech, although they failed on a less frequent relationship (Van Heugten & Johnson, 2010).

These results raise the possibility that adjacent and non-adjacent repetitions are processed differently or follow different developmental paths. This interpretation cannot be excluded on the basis of the existing studies, as they compared adjacent and non-adjacent repetitions (ABB vs. ABA) directly. In such a task, successful discrimination performance can be

achieved by encoding only one of the regularities, e.g. the adjacent one, and discriminating it from the other by exclusion, without detecting its underlying pattern. In other words, infants may detect, say, the adjacent repetition, and treat the other pattern as if it were a random ABC sequence. One possible hypothesis is that adjacent regularities are easier to learn than non-adjacent ones. Indeed, newborns are only able to learn the former and not the latter (Gervain *et al.*, 2008). Also, it has been suggested (Endress *et al.*, 2009; Endress, Dehaene-Lambertz & Mehler, 2007) that the former (but possibly not the latter) are detected by a low-level, automatic repetition- or identity-detector that even some animal species possess (e.g. honeybees can detect and categorize repeating stimuli, even after a delay, as being the same; Giurfa, Zhang, Jenett, Menzel & Srinivasan, 2001). Further, a recent optical imaging study comparing the brain responses of infants aged 0;7 and 0;9 to adjacent repetitions and random controls found greater activation in response to the repetition pattern in the younger age group, and to the random sequence in the older group (Wagner, Fox, Tager-Flusberg & Nelson, 2011). What remains unexplored at the behavioral as well as the neural level is whether at age 0;7, non-adjacent repetitions can also be learned.

The present study addresses this question. In Experiment 1, we test infants' ability to discriminate adjacent and non-adjacent repetition structures against random controls (ABB vs. ABC and ABA vs. ABC). In Experiment 2, we show that this discrimination ability is indeed due to the extraction of structural regularities during familiarization, and not to low-level biases inherent in the stimuli or to spontaneous preferences for repetition structures.

EXPERIMENT 1: INFANTS CAN DISCRIMINATE BOTH ADJACENT AND NON-ADJACENT REGULARITIES FROM RANDOM SEQUENCES

To determine whether infants aged 0;7 are able to learn non-adjacent repetition-based regularities in speech, we tested infants' ability to discriminate ABA structures from otherwise similar ABC controls in a headturn preference paradigm similar to Marcus *et al.*'s (1999) original study. We also tested the discrimination of ABB vs. ABC so that performance on adjacent and non-adjacent repetitions could be compared.

In addition to the usual looking time comparisons averaged across all trials of a given condition, we also tested change in looking times throughout the timecourse of the experiment. In previous imaging work with newborns (Gervain *et al.*, 2008), there is evidence for learning across the experimental session for adjacent repetitions, but not for non-adjacent repetitions. Recently, infants aged 0;7 were also found to show discrimination between

862

ABB and ABC sequences in the final, but not in the initial blocks of an optical imaging study (non-adjacent repetitions were not tested; Wagner *et al.*, 2011). We therefore reasoned that the timecourse of infants' behavior might be informative. Moreover, adjacent and non-adjacent repetitions might elicit different looking patterns over the course of the test session. Such differences might average out across the entire duration of the test phase, but could be detectable with a more fine-grained analysis.

METHOD

Participants

The infants were assigned to one of two conditions. We tested eighteen infants in the ABB condition (11 boys, 7 girls, mean age: 0;7·1, range: 0;6·16–0;7·12). Eleven additional infants were tested, but not included in the data analysis as they failed to complete the test session due to fussiness and crying (9) or have looked for the entire duration of the trial (26 sec) in more than two test trials (2). We tested twenty-two infants in the ABA condition (12 boys, 10 girls, mean age: 0;7·0, range: 0;6·13–0;7·14). Eight additional infants were tested, but were not included in the data analysis as they failed to complete the test session due to fussiness and crying. Infants had no known neurological and developmental disorders. Parents gave written informed consent prior to participation. The study was approved by the ethics board of the University of British Columbia.

Material

Three artificial grammars were created, an adjacent (ABB), a non-adjacent (ABA) and a control (ABC) one. The grammars all generated trisyllabic sequences and used the same consonant–vowel syllable repertoire. Two sets of syllables were used to construct the material: one for familiarization and one for test (see 'Appendix' for the list of syllables used). This requires infants to generalize the underlying pattern during test, rather than rely on syllable frequencies or transitions heard during familiarization. In order to create the familiarization stream, each syllable was combined with all other syllables having a different vowel and a different consonant, ensuring maximal discriminability within words. This resulted in seventy-two different trisyllabic repetition sequences (ABB for one group of infants, ABA for another group), which were concatenated in a random order, separated by silences of 1,000 ms. Syllables were all 270 ms long (consonant: 120 ms, vowel: 150 ms), resulting in trisyllabic sequences of 810 ms, following Gervain *et al.*'s (2008) study. The ABB and ABA familiarization streams were thus approximately 2 min 15 sec long, similarly to Marcus *et al.* (1999). We systematically used the repetition, and not the random grammars, for

familiarization, because they were the target structures we were interested in and because the ABC grammar, not having a clear pattern that could be extracted, was found to be more difficult to learn in newborn studies (Gervain *et al.*, 2008; Gervain *et al.*, 2011). If infants had been familiarized to the ABC random grammar, and had then failed to show discrimination in the test, it would have been impossible to decide whether their null performance had resulted from a failure to learn ABC during familiarization or a failure to discriminate ABC and the repetition grammar in the test. For test, six repetition items and six controls were created (see 'Appendix' for the list of items). The frequency of each syllable across all items and within each position was equated between the two test item types. Test items were 810 ms long. Within a test trial, the same item was repeated fifteen times, separated by silences of 1 sec, resulting in a trial of approximately 26 sec.

The material was synthesized using the French f2 voice of the MBROLA diphone database (Dutoit, 1997). We chose to use a non-native phoneme set, because pilot testing in our laboratory showed that the voices are of better quality with this phoneme set than with the native English voices of MBROLA.

Procedure

We used the headturn preference procedure, similarly to Marcus *et al.* (1999). Infants were tested individually while sitting on a parent's lap in a dimly lit, sound-attenuated cubicle. Parents were listening to music and wearing dark sunglasses to avoid parental influence on infants' behavior. Infants first listened to the familiarization stream, while they watched attention-getting lights at the two sides or the center of the testing cubicle. The blinking of the lights was contingent upon the infants' looking behavior, but not on the sounds, which continued playing irrespectively of whether the infant was looking or not. During the experiment, an experimenter, blind to the stimuli and seated outside the testing cubicle, monitored infants' looking behavior and controlled the lights and the stimuli. Infants were videotaped during the experiment for off-line coding.

Immediately after familiarization, infants were tested in twelve test trials. Half of the trials were repetition trials implementing the regularity the infant was exposed to during familiarization (e.g. ABB for the ABB familiarization group). The other half were control ABC trials, identical for the ABB and ABA groups. Each trial started with the blinking of the central light to attract infants' attention. Once infants attended to the central light, one of the side lights started blinking and the central light was extinguished. During test trials, the presentation of the sound stimuli was contingent upon infant looking. When infants stably fixated on the blinking side light, the associated
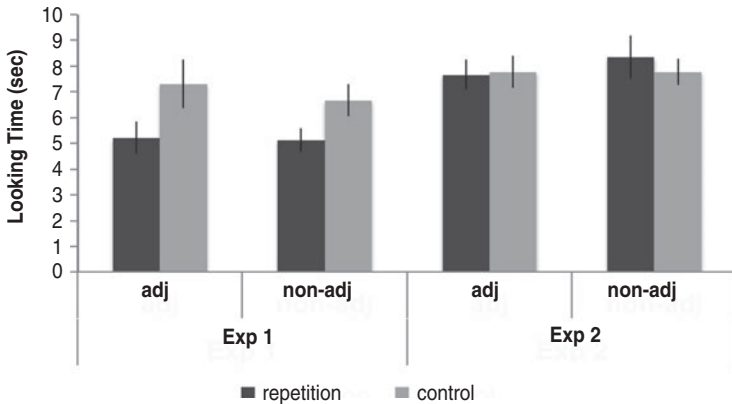
Fig. 1. The average looking times to the repetition (dark gray) and control (light gray) test items in Experiments 1 (left) and 2 (right). Error bars indicate the standard error of the mean.

test item started playing from a loudspeaker on the corresponding side. The sound file continued until the end (26 sec) or until infants looked away for more than 2 sec. After this, a new trial began. The order and side of presentation of the test trials was randomized and counterbalanced across participants in such a way that at most two consecutive trials could be of the same type.

RESULTS

Average looking times to the repetition and control items during test are shown in Figure 1 (left part). We performed a two-way mixed ANOVA with Repetition Type (ABB/ABA) as a between-subjects factor and Test Item Type (repetition/control) as a within-subject factor. We observed a highly significant main effect of Test Item Type $(F(1, 79) = 18 \cdot 358$, $p < 0 \cdot 0001$, partial $\eta^2 = 0 \cdot 3257)$, as infants looked significantly longer at the novel, unfamiliarized control sequences than at the familiar repetition sequences. This effect was present in both familiarization groups. In the ABA group, mean looking time to the repetition items was 5·30 sec, to the control item 6·66 sec (Scheffe post-hoc test: $p = 0 \cdot 016$). In the ABB group, mean looking time to the repetition items was 5·21 sec, to the control item 7·30 sec (Scheffe post-hoc test: $p = 0 \cdot 001$). No other main effect or interaction was significant.

The timecourse of looking behavior across consecutive trial pairs for the two groups is shown in Figure 2 (left part). We conducted a three-way mixed ANOVA with Repetition Type (ABB/ABA) as a between-subjects factor as
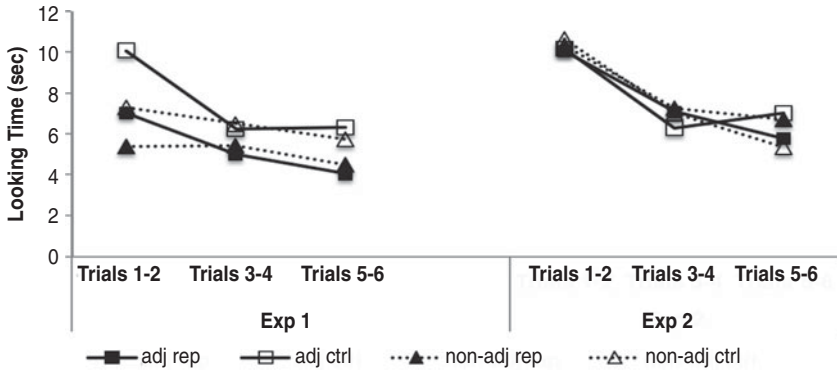
Fig. 2. The average looking times for consecutive trial pairs in Experiments 1 (left) and 2 (right). Continuous lines with square markers indicate looking times for the adjacent repetition group, dotted lines with triangle markers for the non-adjacent repetition group. Filled markers represent looking times to the repetition items, empty markers to the control items.

well as Test Item Type (repetition/control) and Time (Trials 1–2/Trials 3–4/ Trials 5–6)[2] as within-subject factors. We obtained a significant main effect of Time ($F(2, 180) = 7\cdot927$, $p < 0\cdot001$, partial $\eta^2 = 0\cdot0805$), as looking times were significantly longer in Trials 1–2 than in Trials 3–4 (Scheffe post-hoc test: $p = 0\cdot026$) and Trials 5–6 (Scheffe post-hoc test: $p < 0\cdot001$). We also found a highly significant main effect of Test Item Type ($F(1, 181) = 16\cdot679$, $p < 0\cdot0001$, partial $\eta^2 = 0\cdot0843$), as control items induced longer looking times than repetition items. No other main effects or interactions were significant.

DISCUSSION

The results of the grand average ANOVA suggest that infants aged 0;7 are able to discriminate both adjacent and non-adjacent repetitions from random controls after familiarization with the repetition pattern. No differences were found between the two patterns. This was further confirmed by the timecourse analysis, which also revealed successful discrimination in both groups throughout the entire test phase. While looking times were longer overall in the first two trials, the control items reliably induced longer looking times than repetition items in each trial pair for both patterns. This suggests that adjacent and non-adjacent items are processed similarly at this age, at least in the speech domain.

However, as the control items constituted the novel pattern in this experiment, it might be the case that infants paid more attention to them not

[2] We also performed similar ANOVAs with individual trials (Trial 1/Trial 2/Trial 3/Trial 4/Trial 5/Trial 6) and trial triads (Trials 1–3/Trials 4–6) as levels of the Time factor and obtained similar results.

866

because they were novel with respect to the familiarization pattern, but because they were more variable than the repetition patterns, as they had three rather than two different syllables. This still constitutes discrimination, but it is attributable to a mechanism different from pattern extraction or rule learning. Rather, a mechanism more akin to automatic repetition detection (or variability detection – the other side of the same coin) could be involved.

EXPERIMENT 2: DO INFANTS HAVE A SPONTANEOUS PREFERENCE FOR VARIABILITY OR REPETITIONS?

We conducted a second experiment, which was similar to Experiment 1, except that infants were given no familiarization. If the preference for the more variable ABC items is spontaneous, then the results of this experiment should parallel those of Experiment 1. If no such preference can be observed, then familiarization can be concluded to have played a role in the results found in Experiment 1. Alternatively, it might be the case that without familiarization, a spontaneous preference for the repetition items, or at least for the adjacent repetitions, emerges. Indeed, it has been suggested that adjacent repetitions, although possibly not distant ones, are processed by an automatic perceptual repetition (Endress *et al.*, 2009). It is therefore possible that infants might show a spontaneous preference for these easily detectable structures over random sequences. Such preference would show that even without training, infants can discriminate sequences with adjacent repetitions from random controls. The same preference, however, might not be present for non-adjacent repetitions.

Examining the timecourse of infants' looking behavior seemed particularly relevant here, as, without familiarization, spontaneous preferences or processing advantages could emerge even more easily, but might average out over the full duration of the experiment. If, for instance, an automatic repetition-detection mechanism was in place, we would predict a fast initial discrimination between controls and the type of repetition items (adjacent and/or non-adjacent) that the mechanism applies to.

METHOD

Participants

The infants were assigned to one of two conditions. We tested eighteen infants in the ABB condition (10 boys, 8 girls, mean age: 0;7·1, range: 0;6·17–0;7·12). Four additional infants were tested, but were not included in the data analysis as they failed to complete the testing session due to fussiness and crying. We tested eighteen infants in the ABA condition (7 boys, 11 girls, mean age: 0;7·0, range: 0;6·10–0;7·10). Four additional infants were tested, but not included in the data analysis as they failed to complete the testing

867

session due to fussiness and crying. Infants had no known neurological and developmental disorders. Parents gave written informed consent prior to participation. The study was approved by the ethics board of the university where the infants were tested.

Material

The material used in this experiment was the same as the test items in Experiment 1. No familiarization was used.

Procedure

The procedure used in this experiment was identical to Experiment 1, except for the absence of familiarization.

RESULTS

Looking times averaged across trials to the repetition and control items during testing are shown in Figure 1 (right part). We performed a two-way mixed ANOVA with Repetition Type (ABB/ABA) as a between-subjects factor and Test Item Type (repetition/control) as a within-subject factor. We found no significant main effects or interactions (Repetition Type: $F(1, 71) = 0.201$, n.s.; Test Item Type: $F(1, 71) = 0.180$, n.s.; Repetition Type × Test Item Type $F(1, 71) = 0.411$, n.s.). Post-hoc tests were all non-significant.

We also performed a timecourse analysis as for Experiment 1. Looking times for consecutive trial pairs are shown in Figure 2 (right part). The three-way mixed ANOVA with Repetition Type (ABB/ABA) as a between-subjects factor as well as Test Item Type (repetition/control) and Time (Trials 1–2/Trials 3–4/Trials 5–6)[3] as within-subject factors revealed a highly significant main effect of Time ($F(2, 169) = 23.872$, $p < 0.0001$, partial $\eta^2 = 0.2192$), as looking times were longer during Trials 1–2 than during Trials 3–4 (Scheffe post-hoc test: $p < 0.0001$) and Trials 5–6 (Scheffe post-hoc test: $p < 0.0001$).

To compare the two experiments, we also performed a grand ANOVA, pooling together data from Experiments 1 and 2. This three-way mixed ANOVA with Familiarization (present = Exp1/absent = Exp2) and Repetition Type (ABB/ABA) as between-subjects factors and Test Item Type (repetition/control) as a within-subject factor revealed a main effect of Familiarization ($F(1, 151) = 10.150$, $p = 0.002$, partial $\eta^2 = 0.2841$), as infants

[3] We also performed similar ANOVAs with individual trials (Trial 1/Trial 2/Trial 3/Trial 4/Trial 5/Trial 6) and trial triads (Trials 1–3/Trials 4–6) as levels of the Time factor and obtained similar results.

868

in Experiment 2, having received no familiarization, showed longer looking times overall than infants in Experiment 1. There was also a main effect of Test Item Type ($F(1, 151) = 5 \cdot 138$, $p = 0 \cdot 026$, partial $\eta^2 = 0 \cdot 0666$), as the control items provoked longer looking times than the repetition items. In addition, there was a significant interaction between Familiarization and Test Item Type ($F(1, 151) = 8 \cdot 723$, $p = 0 \cdot 004$), due to longer looking times to control than to repetition items in Experiment 1 (control > repetition, Scheffe post-hoc test: $p = 0 \cdot 0003$), but not in Experiment 2 (repetition vs. control, n.s.).

DISCUSSION

The results of Experiment 2 indicate that, without familiarization, infants have no spontaneous preference for either test item type. This suggests that the longer looking times observed for control items in Experiment 1 are not due to the more variable nature of these items. The results of the timecourse analysis also point in this direction, as no difference between repetition items and controls was found at any point during the test. Rather, the longer looking times for control items seen in Experiment 1 arose because infants were able to extract the underlying structure of the repetition grammar they were familiarized with and discriminated it from the random sequence of the control items. The significant interaction in the grand ANOVA further confirmed this result.

A second important result of Experiment 2 is the absence of any difference between the adjacent and non-adjacent repetition groups. The similarity between the two patterns was also confirmed by the timecourse analysis and the grand AVONA, as repetition type had no main effect, nor did it enter into significant interactions.

The grand ANOVA also revealed a main effect of familiarization on looking times. Infants in Experiment 1 showed shorter looking times overall in the test than infants in Experiment 2. This is expected, as they received a 2-minute-long familiarization, leading to increased fatigue during test. The main effect of Test Item Type in the grand ANOVA was mainly driven by the shorter looking times to the repetition items in Experiment 1.

GENERAL DISCUSSION

The results from these two experiments provide the strongest evidence to date that infants aged 0;7 are able to learn non-adjacent repetition in the stream of speech. Importantly, longer looking to the control items in Experiment 1 in the adjacent as well as non-adjacent condition is not simply due to the more variable nature of these items. If such a preference had existed, it should have also been observed in Experiment 2, contrary to fact.

869

Rather, these results indicate that discrimination in Experiment 1 arose because infants were able to extract the underlying structure of the repetition grammars they were familiarized with and discriminate it from the control items.

While we argue that infants aged 0;7 can learn both adjacent and non-adjacent repetitions, we do not deny that adjacent repetitions might be easier to process under some circumstances. Johnson *et al.* (2009), for instance, found that even infants aged 0;11 were unable to learn ABA patterns (contrasted in the test with adjacent repetition patterns), when shown sequentially presented visual stimuli, but succeeded with ABB and AAB structures (contrasted with ABA). These and similar asymmetries reported in the literature may be related to the increased cognitive load of the stimulus presentation mode (sequential, as opposed to simultaneous) or the modality (visual, as opposed to speech). Indeed, Saffran, Pollak, Seibel and Shkolnik (2007) found that when the visual stimuli used were presented simultaneously and were natural (breeds of dogs) as opposed to geometric figures, even infants aged 0;7 succeeded (at least in the usual adjacent vs. non-adjacent repetitions comparison). Also, some modalities might serves as better input for the pattern extraction mechanism than others. Marcus *et al.* (2007) reported that infants were only able to perform the ABB vs. ABA discrimination on sequences implemented as animal sounds, musical tones and timbre if they were first trained on speech sequences, i.e. they were able to transfer the extracted structure. Later, Frank *et al.* (2009) showed that infants as young as 0;5 were able to learn adjacent and non-adjacent repetition sequences when they were presented multimodally, i.e. as a combination of visual and speech stimuli, but not when only one modality was used. It needs to be noted that all these studies familiarized infants with one of the two repetition types and tested them on the discrimination of the two types. This procedure is arguably more demanding than the repetition vs. random control test we used in our experiments in that the ABA vs. ABC discrimination necessarily requires non-adjacent elements to be recognized as identical, while the ABA vs. ABB comparison can be performed on the basis of strictly local relations, i.e. by recognizing the adjacent, but not the non-adjacent repetition.

Our results only provide a snapshot of the development of non-adjacent regularity learning during the first year of life. More research will be needed to explore the change that takes place between birth and age 0;7, allowing infants to start tracking non-adjacent regularities. At least two explanations are possible. It might be the case that adjacent and non-adjacent repetitions are processed by two different mechanisms, which also have different developmental trajectories. The mechanism responsible for adjacent repetitions is operational at birth, whereas the one processing non-adjacent repetitions appears later in development. One candidate mechanism that has

been proposed to account for adjacent repetitions is an automatic, low-level repetition or identity detector (Endress *et al.*, 2009). As this mechanism is also present in non-linguistic animals, e.g. honeybees (Giurfa *et al.*, 2001), it is not implausible to assume that it might be present in humans as early as birth. By contrast, non-adjacent repetitions might be processed by a more abstract, symbolic mechanism, which requires time to develop. Indeed, Gervain *et al.* (2008) provided evidence that newborn humans can detect adjacent but not non-adjacent repetitions. Alternatively, it might be the case that the same mechanism is involved in the processing of both adjacent and non-adjacent regularities, but their scope is restricted at birth due to memory or attentional limitations. Our current results need to be interpreted with caution in this context, given that in the headturn paradigm, a lack of preference does not entail a lack of discrimination. However, our experiments seem to show no difference between adjacent and non-adjacent repetitions either with or without familiarization. Crucially, in the latter context, no spontaneous preference arises. This provides at least some evidence against a dual-mechanism account. The presence of an automatic repetition-detector cannot be excluded, but we can at least conclude that infants aged 0;7 seem not to use it in these tasks to develop a preference for (adjacent or non-adjacent) repetitions.

CONCLUSION

Our experiments provide evidence that infants aged 0;7, i.e. prelinguistic infants, are able to extract non-adjacent repetition-based regularities from speech similarly to adjacent ones. This finding contributes to a better understanding of the mechanisms underlying language acquisition, as non-adjacent dependencies are fundamental to language structure.

REFERENCES

Dutoit, T. (1997). *An introduction to text-to-speech synthesis*,Vol. 3. Dordrecht, Boston: Kluwer Academic Publishers.

Endress, A. D., Dehaene-Lambertz, G. & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition* **105**, 577–614.

Endress, A. D., Nespor, M. & Mehler, J. (2009). Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences* **13**, 348–53.

Frank, Michael C., Slemmer, Jonathan A., Marcus, Gary F. & Johnson, S. P. (2009). Information from multiple modalities helps five-month-olds learn abstract rules. *Developmental Science* **12**, 504–509.

Gervain, J., Macagno, F., Cogoi, S., Pena, M. & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 14222–27.

Gervain, J., Mehler, J., Werker, J. F., Nelson, C. A., Csibra, G., Lloyd-Fox, S., Shukla, M. *et al*. (2011). Near-infrared spectroscopy: A report from the McDonnell infant methodology consortium. *Developmental Cognitive Neuroscience* **1**, 22–46.

Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature* **410**, 930–33.

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

Gómez, R. L. & Maye, J. (2005). The developmental trajectory of nonadjacent dependency learning. *Infancy* **7**, 183–206.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. Cambridge, MA: MIT Press.

Johnson, S. P, Fernandes, K. J., Frank, M. C, Kirkham, N., Marcus, G., Rabagliati, H. & Slemmer, J. A. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy* **14**, 2–18.

Lany, J. & Gómez, R. L. (2008). Twelve-month-old infants benefit from prior experience in statistical learning **19**, 1247–52.

Marcus, G. F, Fernandes, K. J. & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science* **18**, 387–91.

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

Saffran, J. R., Pollak, S. D., Seibel, R. L. & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition* **105**, 669–80.

Santelmann, L. M. & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition* **69**, 105–134.

Van Heugten, M. & Johnson, E. K. (2010). Linking infants' distributional learning abilities to natural language acquisition. *Journal of Memory and Language* **63**, 197–209.

Wagner, J. B., Fox, S. E., Tager-Flusberg, H. & Nelson, Charles A. (2011). Neural processing of repetition and non-repetition grammars in 7- and 9-month-old infants. *Frontiers in Psychology* **2**.

APPENDIX

Syllables used for familiarization:
/ba/, /bi/, /bɔ/, /la/, /li/, /lɔ/, /ma/, /mi/, /mɔ/, /ʃa/, /ʃi/, /ʃɔ/

Syllables used for test:
/pe/, /pu/, /pə/, /ne/, /nu/, /nə/, /fe/, /fu/, /fə/

ABB test items:
/pefəfə/, /pefufu/, /nepəpə/, /nepupu/, /fenunu/, /fenənə/

ABA test items:
/fəpefə/, /fupefu/, /pənepə/, /punepu/, /nufenu/, /nəfenə/

ABC test items:
/penufə/, /nefupə/, /fenupə/, /nupəfe/, /pufəne/, /nunəpe/

872