

# Auditory perception at the root of language learning

Jutta L. Mueller, Angela D. Friederici, and Claudia Männel<sup>1</sup>

Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany

Edited by Michael Merzenich, W. M. Keck Center for Integrative Neuroscience, San Francisco, CA, and approved July 26, 2012 (received for review March 13, 2012)

**Learning a spoken language presupposes efficient auditory functions. In the present event-related potential study, we tested whether and how basic auditory processes are related to online learning of a linguistic rule in infants and adults. Participants listened to frequent standard stimuli, which were interspersed with infrequent pitch deviants and rule deviants, violating a non-adjacent dependency between two syllables. Only infants who showed the more mature mismatch response for the pitch deviants (i.e., a negativity) showed a mismatch response to the rule deviants. Concordantly, the small group of adults who showed evidence of rule learning showed larger mismatch effects for pitch processing. We conclude that the ability to extract linguistic rules develops in early infancy and is tightly linked to functional aspects of basic auditory mechanisms.**

language acquisition | AXB rules | pitch perception | mismatch negativity

Human language is based on an acoustically transmitted signal. The successful language learner needs to decode the linguistic content of a complex auditory signal into its component units and their relation to each other, thus deriving words and rules. Although sufficient speech input is widely considered to be crucial for language learning, the perceptual abilities that form the gateway to spoken language have long been neglected in first and second language acquisition research. Basic auditory perception may, however, be an important determinant of language learning processes across the range of normal and abnormal development.

There is empirical evidence supporting the idea that early auditory abilities impact on later outcomes of language development in normal infants and populations with language-related disorders (1–3). Furthermore, auditory brainstem responses in language-impaired children suggest that low-level auditory processes contribute to the pathogenesis of language disorders (4). In adults, individual differences in perceptual abilities correlate with language-processing abilities in their native and a second language (5, 6). The findings suggest a potential causal relationship between basic auditory processing ability and the efficiency of language learning in infancy and adulthood.

Although infants do not produce complex language in their first year of life, they show remarkable abilities to extract regular patterns from speech input early on. Eight-month-old infants, and even newborns, have been shown to be sensitive to transitional probabilities between syllables defining word-like units (7, 8). Syllable repetitions can be detected from birth (9) and dependencies between nonadjacent units of speech can be detected as early as 4 mo of age (10). Learners seem to be able to exploit various distributional and acoustic cues to detect words and rules in speech input. Infants, for example, take advantage of prosodic cues for detecting possible words in the linguistic input (11). Similarly, prosodic cues seem to assist adults' extraction of grammatical patterns from speech input (12, 13).

Despite the importance of acoustic cues for language learning tasks, and despite the proven impact of auditory perception on language development, empirical evidence of a synchronistic impact of auditory perceptual abilities on online language learning is lacking. In this study, we demonstrate that basic auditory discrimination skills are linked to the learning of a simple grammatical rule in a population of healthy infants and adults.

We chose pitch perception as a test for auditory perceptual function. Frequency, which is the main carrier of pitch, is a distinctive sound feature representing resonance properties of the vocal tract, and thus information about differences in sound production. The importance of frequency information for language comprehension becomes evident when the frequency spectrum is degraded and leads to severe problems in speech recognition (14). On the other hand, an enhanced frequency spectrum may boost language learning, as it was shown to be beneficial for the discrimination of vowel categories by infants at about 6 mo of age (15).

We chose nonadjacent dependency rules as a prototypical test for grammatical rule learning. Complex syntactic structures afford building relations between distant parts of a sentence; for example, between the noun “boy” and the third person suffix “-s” in the sentence “The little boy who lives in our neighborhood always smiles.” Learners have to keep track of nonadjacent dependencies to decode such structures. For a model mimicking such rules, we focused on so-called AXB structures, in which A predicts B with an intervening element X. Both adults, and infants starting from 4 mo of age, can learn AXB structures by merely listening to correct examples (10, 16, 17), although girls seem to show a small advantage compared with boys at the age of 12 mo (18).

To examine auditory perception and rule learning in a behavior-independent manner, we applied the auditory oddball paradigm, in which infrequent deviant stimuli are presented among a series of standard stimuli. Electrophysiologically, deviants elicit a mismatch response (MMR). In adults, this MMR appears as a mismatch negativity (MMN), which is widely accepted as indicating cortical processes of memory-based auditory change detection (19). The MMN has been observed for both simple auditory discrimination, as well as relatively complex and even abstract regularities (20). Infants' MMNs can be measured from birth (21, 22) and show a specific developmental pattern, beginning with a positive response in early infancy to a more mature negative-going MMN later on (23–25). Notably, MMN development is feature-specific; that is, an adult-like negativity may occur at different developmental times across various auditory features (23). These properties make the oddball paradigm ideally suited to the study of the relation between auditory perception and rule learning, particularly because it can be used with infants as well as adults. The present study uses pitch- and rule-related MMNs to investigate the relation between auditory perception and the discovery of rule-based dependencies in speech across development. To investigate interindividual differences in the ability to detect rule-based dependencies, we make use of the maturational transition from a negative toward a positive MMR in infancy and of behavioral measures of rule learning in adulthood.

Author contributions: J.L.M. designed research; J.L.M. and C.M. performed research; J.L.M. and C.M. analyzed data; and J.L.M., A.D.F., and C.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

<sup>1</sup>To whom correspondence should be addressed. E-mail: muellerj@cbs.mpg.de.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1204319109/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1204319109/-DCSupplemental).

The stimuli were sequences of naturally spoken syllables recorded by a female speaker. Standard stimuli were triplets of consonant-vowel syllables of the form AXB, in which the first syllable A predicted the third syllable B. The intervening syllable X varied between 20 different exemplars. We used the two standard A...B frames *fi...to* and *le...bu*, with the intervening X syllables *ka, we, mi, no, gu, sa, me, ri, ro, ku, ma, ke, gi, ko, su, wa, xe, ki, so*, and *mu*. Rule deviants contained a violation of the remote dependency between the first and the third syllable (*fi...bu* and *le...to*) and pitch deviants had an increased pitch of ~11% compared with the average across all syllables. This magnitude of pitch change is well above a normal perceptual threshold and infants have shown robust MMRs, even at lower magnitudes of pitch change (24). Participants listened to a stimulus stream in which pitch and rule deviants were randomly interspersed in a series of standard stimuli (Fig. 1).

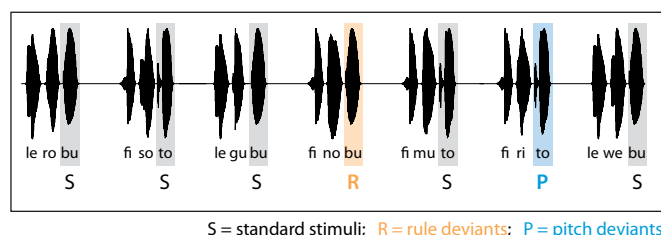
As the present study takes a developmental perspective, infant and adult learners were included. Because previous findings suggest sex differences specifically for verbal learning tasks (18, 26), we tested a comparable number of males and females in all participant groups, allowing us to look at sex-specific learning effects. The infant group (Exp. 1) consisted of 3-mo-old infants who listened to the stimuli while being held by a parent. During early language development, speech stimuli automatically attract infants' attention (27). In the case of adult learners, however, it is not known whether learning can take place in the absence of specific attention and task participation because previous studies have all used attentive exposure conditions (16, 28). To address this issue, we measured two groups of adults. One adult group (Exp. 2) was tested under passive listening conditions, similar to the infant group. The second adult group (Exp. 3) comprised adult participants who were given a dual target detection task, requiring a button press in response to deviants that violated either a "regularity with respect to pitch" or a "regularity with respect to the correct order of the syllables."

## Results

**Experiment 1.** Fig. 2 shows infants' event-related potential (ERP) responses for the pitch and the rule conditions for four different groups, with the ERP response pattern to pitch deviants and sex as the group-defining factors.

All infants were categorized according to the maturational status of their MMR (negative vs. positive polarity) to pitch deviants vs. standards, with the positivity reflecting a less-mature response than the negativity. Thirty-two infants displayed a mature negative MMR (neg MMR, 15 girls). Thirty-three infants showed the less mature positive MMR (pos MMR, 18 girls).

First, ERP responses to pitch deviants were statistically analyzed in the time window (TW) that was chosen to categorize the neg MMR and the pos MMR groups (60–260 ms). ANOVAs and step-down analyses revealed broadly distributed effects of PITCH for both the neg MMR and the pos MMR group (Table 1). Second, an ANOVA was conducted for rule deviants between



**Fig. 1.** Oscillogram of a series of standard and deviant stimuli. Single syllables in rule deviants (R) are acoustically identical to standard stimuli (S) in contrast to pitch deviants (P).

60 and 260 ms relative to the onset of the final syllable. Results revealed a significant interaction of RULE  $\times$  MMR  $\times$  SEX (Table 1). Step-down analyses showed that only infants in the neg MMR groups showed a significant RULE effect, however, with opposite polarities of the amplitudes in boys and girls (boys: 1.89  $\mu$ V; girls:  $-2.58 \mu$ V). In contrast, the pos MMR groups did not show significant effects. Thus, only those infants who showed a negative MMR for pitch deviants showed evidence of rule learning.

**Experiment 2. Behavioral results.** For adults who were tested under passive listening conditions, there was no evidence of successful rule learning in the subsequent familiarity judgment task (correct answers: mean 53.1%, SD 5.9%).

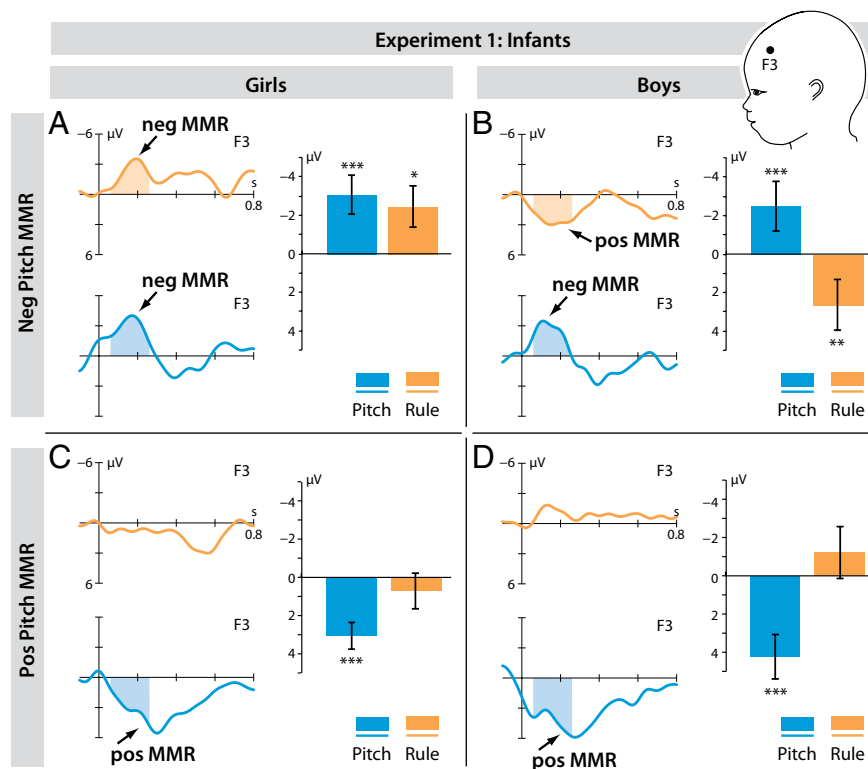
**ERP results.** Fig. 3 shows adults' ERP responses for the pitch and the rule conditions. For the pitch condition, there was a significant main effect of PITCH ( $F_{1,18} = 7.89, P < 0.05$ ) and an interaction of PITCH  $\times$  REGION ( $F_{1,18} = 9.69, P < 0.01$ ) between 120 and 280 ms. This result was because of a negativity (MMN) that was only present over anterior ( $F_{1,18} = 11.16, P < 0.01$ ) and central electrode sites ( $F_{1,18} = 7.30, P < 0.05$ ). In addition, there was a subsequent, long-lasting anterior negativity between 480 and 800 ms reflected in a PITCH  $\times$  REGION interaction ( $F_{1,18} = 18.36, P < 0.001$ ) driven by a simple main effect of PITCH over anterior electrode sites ( $F_{1,18} = 12.63, P < 0.01$ ). Thus, statistical analyses revealed an early and a late effect for the pitch condition over anterior electrode sites. There were no significant effects including the factor RULE in the ANOVA and no linear relationship between the effects in the pitch condition and the rule condition as tested in an additional regression analysis (*SI Regression Analysis*).

**Experiment 3. Behavioral results.** For adults who were tested under attentive exposure conditions, both the target detection task and the subsequent familiarity judgment task revealed that the participants were clearly clustered in two groups of performers: namely, rule learners (10 participants, detection rate: 45.6%, SD 18.78; familiarity judgment: 97.75% correct, SD 2.4) and nonlearners (26 participants, detection rate: 9.39%, SD: 10.44; familiarity judgment: 53.7% correct, SD 4.6). The mean target detection rate for the pitch deviants was not statistically different between groups (learners: mean 69.2%, SD 11.7; nonlearners: mean 71.2%, SD 9.9).

**ERP results.** The behavioral categorization enabled a direct test of whether rule learners and nonlearners process rule and pitch deviants differently. Only ERP data of trials of detected pitch and rule deviants were analyzed, with the exception that, for the rule condition in nonlearners, all trials (because of the absence of detected targets) were included. Fig. 4 shows adults' ERP responses for the pitch and the rule conditions separately for learners and nonlearners.

We report only the effects that include both factors GROUP and PITCH. Between 140 and 380 ms, there was a significant GROUP  $\times$  PITCH interaction ( $F_{1,34} = 7.44, P = 0.01$ ). This result was because of a larger amplitude of the negative effect (MMN/N2) in the learner group ( $-2.37 \mu$ V) compared with the nonlearner group ( $-1.28 \mu$ V). In the TW from 640 to 720 ms, there was a significant interaction of GROUP  $\times$  PITCH  $\times$  LATERALITY  $\times$  REGION ( $F_{8,271} = 2.80, P < 0.05$ ). Further tests revealed that the interaction was driven by GROUP  $\times$  PITCH interactions over left posterior, left medial posterior, right medial posterior, and right posterior electrode sites (all  $F_{s1,34} > 5.0$ , all  $P < 0.05$ ). This result was because of larger positivities in the learners compared with nonlearners (e.g., left posterior region of interest; learners: 5.62  $\mu$ V, nonlearners: 3.16  $\mu$ V).

For the rule deviants, there were GROUP  $\times$  RULE interactions starting from 200 ms after stimulus onset. The results



**Fig. 2.** Rule learning is linked to pitch processing in infants. ERP difference waveforms and bar plots representing mean amplitudes (deviants – standards) in the significant TW 60–260 ms at the representative electrode F3, which contributed to the significant main effects across comparisons. Significant effects are marked with shaded areas in the waveforms. In the bar plots, significant effects are marked with asterisks ( $*P \leq 0.05$ ,  $**P \leq 0.01$ ,  $***P \leq 0.001$ ). (A) Girls with neg MMR in the pitch condition show negativity for rule condition. (B) Boys with neg MMR in the pitch condition show positivity for rule condition. Girls (C) and boys (D) with pos MMR in the pitch condition do not show any significant effects for the rule condition.

showed significant effects including RULE in TWs from 200 to 1,000 ms poststimulus onset for learners only (Table 2).

In sum, learners and nonlearners differed in their ERP patterns for pitch discrimination (MMN/N2, P3) and for rule discrimination (N2/P3), with learners showing enhanced amplitudes for all ERP components. An additional regression analysis revealed that the pitch-related MMN/N2 predicted the rule-related P3 across learners and nonlearners (*SI Regression Analysis* and Fig. S1).

## Discussion

The present findings show a clear relation between an electrophysiological measure of pitch discrimination and the ability of infants, and also adults, to extract a linguistic rule.

**Table 1. Infants: Significant results of omnibus and step-down ANOVAs for pitch and rule conditions**

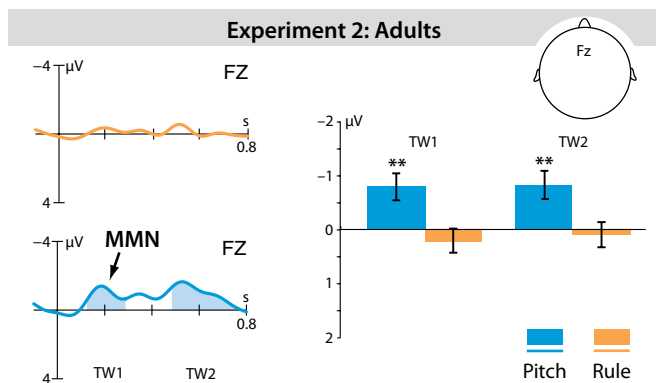
Pitch condition	df	F	Rule condition	df	F
P × MMR	1,61	40.84***	R × MMR × Sex	1,61	9.76**
P × Reg × MMR	1,61	11.73***	Neg MMR-Boys: R	1,16	7.85**
Neg MMR: P	1,31	12.43***	Neg MMR-Girls: R	1,14	6.66*
Pos MMR: P × Reg	2,64	8.75**			
Ant	1,9	27.74***			
Cen	1,9	46.90***			
Post	2,18	9.35**			

Step-down analyses are given after the superordinate analysis. Ant, anterior; Cen, central; P, pitch; Post, posterior; R, rule; Reg, region.

\* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ .

The experiment with 3-mo-old infants shows that those infants who display a negative MMR to the pitch deviants successfully extracted the rule-based dependency from the auditory input, as reflected in the amplitude difference between ERPs for rule deviants vs. standards. Previous studies testing infants' learning of nonadjacent dependency rules behaviorally did not observe any learning before 12 mo of age (17, 18). However, a recent ERP study indicated that infants might be sensitive to nonadjacent dependency rules in a novel language already at the age of 4 mo (10). The present finding with 3-mo-olds further corroborates this finding. More importantly, the present results show that the ability to detect rule-based dependencies is related to the polarity of the observed MMRs in response to pitch discrimination and, thus, on the maturational status of auditory perception. The exact functional interpretation of the positive and negative MMRs in early infancy is still a topic of debate. However, it has been argued that the negative MMR in infants reflects mature, adult-like memory-based deviant detection, and the positive MMR might reflect more general processes related to neural adaptation (23) or different states of alertness and attention (25, 29). Notwithstanding these arguments, it is clearly evident from the literature that the positivity occurs earlier than the negativity on the developmental timeline (23, 24), and thus a maturational interpretation of the present findings seems warranted. The current results suggest that there is either a causal link between the ability of frequency discrimination and grammatical rule learning or a parallel development in both domains. A direct causal link could operate via the role frequency information plays in phoneme discrimination (15, 30), which is a precondition for detecting long-range dependencies between phonemes. Alternatively, both pitch perception and rule





**Fig. 3.** No rule learning under passive listening conditions in adults. ERP difference waveforms and bar plots representing mean amplitudes (deviants – standards) at the representative electrode Fz, which contributed to the significant effects in the two TWs [a negativity in TW1: 120–280 ms (MMN), and a negativity in TW2: 480–800 ms]. Significant effects are marked with shaded areas in the waveforms. In the bar plots, significant effects are marked with asterisks (\*\* $P \leq 0.01$ ). Participants show an MMN and a late negativity in the pitch condition and no significant effects in the rule condition.

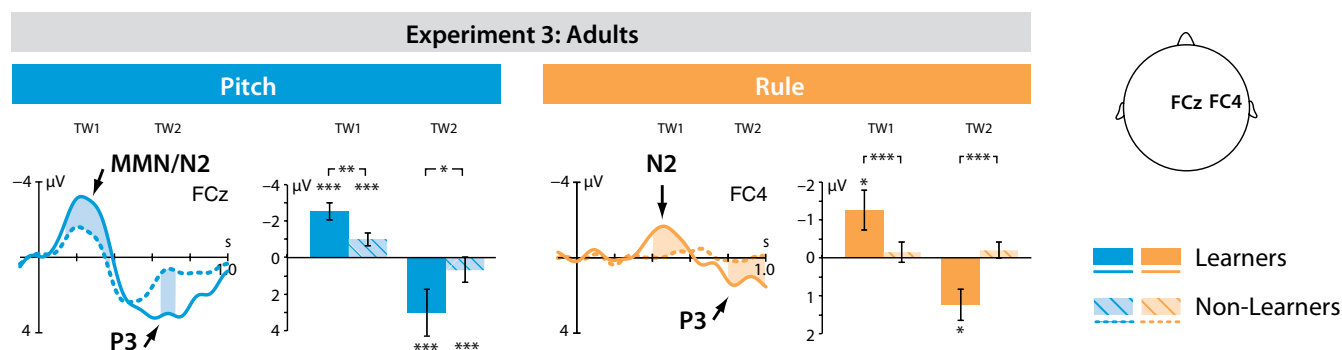
learning could be influenced by a third factor, such as lower efficiency of auditory sensory memory or even a domain-general cognitive mechanism. Basic parameters of physiological maturation as explanatory variables, however, can be excluded, as there were no significant differences between the four infant groups with respect to age, gestational age, and birth weight (Table S1). What is clear from the present study, however, is that a more mature MMR response to pitch discrimination co-occurs with a better grammatical rule learning ability.

Over and above the link between auditory perception and grammatical rule learning, we found sex differences in the polarity of the MMR in infants. Although the mere presence of a mismatch effect evidences the fact that the rule-based dependency must have been extracted, the differential responses indicate different developmental stages, with girls being on a more advanced stage than boys. According to the functional interpretation of He, Hotson, and Trainor (24), the positive MMR reflects a refractoriness-based process and the negative MMR reflects memory-based discrimination. If this interpretation is correct, the rule-learning effects for boys and girls are based on different underlying processes. Neural adaptation as an account

for boys' rule learning in our paradigm is plausible if the auditory system works in a predictive manner, with the final syllable receiving preactivation before its occurrence. Although we cannot prove this within our study, there is independent evidence that prediction occurs in auditory sequence processing (31). With respect to our data, this finding would imply that boys' rule learning is indicated by an enhanced neural response to the nonpredicted syllable at a precognitive level, but girls' rule learning takes place at a more cognitive, memory-based level. Regardless of the nature of the underlying neural mechanisms, developmental differences in the polarity of the MMR across sound features have been observed previously (23), and so have sex differences in early sound discrimination (32). In the light of these findings, it is not surprising that the polarities of the pitch MMR and the rule MMR differ across sexes. The observed differences between boys and girls could be mediated by the hormone testosterone, which has been found to negatively impact phonological discrimination abilities in 1-mo-olds (32) and are in line with evidence showing that girls outperform boys in remote dependency learning at later developmental stages (i.e., at 12 mo) (18) and in general verbal abilities during childhood (26, 34).

Interestingly, the adults in our study who were tested under the same conditions as the infants, namely under passive exposure, did not show any evidence of rule learning. Only when an explicit task was introduced did a rule-learning effect emerge, and even then, it only occurred in 26% of the participants. Although learners and nonlearners did not differ behaviorally in their target detection rate for the pitch condition, they demonstrated differences in their ERP pattern. Compared with nonlearners, learners showed an enhanced amplitude with respect to the MMN/N2 component and the later P3 in the pitch condition. Thus, participants who behaviorally and consciously detect a rule deviant of a remote dependency between syllables, showed enhanced responses in the basic auditory discrimination task. Importantly, auditory discrimination responses (MMN/N2) in turn predict the electrophysiological indicator of rule learning (P3). This direct link between rule-learning ability and differences in pitch processing in adults validates the finding of a strong relation between auditory pitch discrimination and rule learning for infants, and generally shows that successful rule learners process pitch information differently from nonlearners.

Adult rule learners showed a bipolar pattern of an anterior N2 and a P3 in response to detected rule deviants. Both ERP components have previously been reported for sequence learning tasks and indicate deviance processing mechanisms in the pres-



**Fig. 4.** Rule learning is linked to pitch processing in adults during active listening. ERP difference waveforms and bar plots representing mean amplitudes (deviants – standards) plotted for learners and nonlearners separately at representative electrodes, which contributed to the significant pitch and rule effects. Significant differences between learner groups are marked with shaded areas in the waveforms. In the bar plots, significant effects are marked with asterisks (\* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ ). In the pitch condition, shown at the representative electrode FCz, significant differences were present in two TWs [a negativity in TW1: 140–380 ms (MMN/N2), and a positivity in TW2: 640–720 ms (P3)]. Similarly, in the rule condition, shown at the representative electrode FC4, significant differences were present in two TWs [a negativity in TW1: 400–600 ms (N2), and a positivity in TW2: 800–1,000 ms (P3)]. Learners show enhanced responses in the pitch condition and in the rule condition compared with nonlearners.

**Table 2. Adult learners under active exposure: Significant results of omnibus and step-down ANOVAs for the rule condition**

TW (ms)	Rule condition	df	F
200–300	R × Lat × Reg	8,72	4.41**
320–400	R × Lat × Reg	8,72	3.18*
400–600	Right ant	1,9	7.19*
	R × Reg	2,18	16.51**
	Ant	1,9	7.37*
600–800	Post	1,9	5.98*
	R × Reg	2,18	18.15***
	R × Lat × Reg	8,72	2.64*
	Left: cen, med post, post	1,9	11.6–15.62**
800–1,000	Left: med cen; Right: cen, med post, post	1,9	6.76–8.52*
	R	1,9	6.98*
	R × Lat × Reg	8,72	6.47***
	Left: med ant, cen, med cen, post; Medial: ant; Right: ant, med ant, cen, med cen, post	1,9	5.5–10.08*

Step-down analyses are given after the superordinate analysis. Ant, anterior; Cen, central; Lat, laterality; Med, medial; Post, posterior; R, rule; Reg, region.

\* $P \leq 0.05$ ; \*\* $P \leq 0.01$ ; \*\*\* $P \leq 0.001$ .

ence of explicit sequential knowledge (33, 35). Notably, the adult N2 effect occurred in a later TW compared with the rule-related MMR in infants. We take this temporal difference as indication of higher automaticity in infants' rule processing compared with adults'. The result that adults only learn the rule-based dependency in the presence of an active task is in line with this interpretation. It is even more striking that only a minority of the adults was able to detect the rule, despite its rather low level of complexity. This finding corresponds to earlier observations that adult language learners generally appear to have difficulty learning rules of novel languages: only a minority is ultimately successful (36). Thus, the present results might indicate a particular aptitude of rule extraction in a subgroup of adult learners that goes hand in hand with enhanced abilities in basic auditory perception.

Taken together, our ERP experiments indicate a strong relation between auditory discrimination abilities and rule learning at both a developmental and an interindividual level. Adults' failure to learn the rule under the same exposure conditions as the infants indicates a developmental loss of automaticity in rule learning under passive listening conditions. In both infants and adults, interindividual differences with respect to basic auditory processing are linked to linguistic rule learning. The present findings lead to an entire new set of questions with regard to the long-term effects of early language learning, its perceptual roots, and possibilities for enhancement through training.

## Methods

**Participants.** The study was approved by the Ethics Committee of the University of Leipzig and conforms to the guidelines of the Declaration of Helsinki (2008). Before the experiment, participants, or their accompanying caregiver, gave written informed consent. All participants were of normal hearing. Adult participants were right-handed and no participant had any history of neurological disorder.

In Exp. 1, 108 infants were invited for testing. Data from 43 infants could not be obtained or analyzed because of crying during EEG preparation or measurement, sleeping for more than 12 min, or high artifact rate in the EEG data. Participants who entered the final analysis were not significantly different with respect to age, gestational age, and birth weight across groups (Table S1). In Exp. 2, 20 volunteers (10 females) participated (mean age = 25.3 y, SD = 2.4). In Exp. 3, 41 volunteers participated. Five datasets had to be excluded because of high artifact rate or insufficient performance in the target detection task (fewer than 50% of hits in the pitch condition or less than 10 trials in the rule condition, despite learning). Nineteen females and 17 males entered final data analysis (mean age = 24.8 y, SD = 2.5).

Ten participants were assigned to the learner group because they detected more than 10 rule deviants (12.5%) during the ERP experiment and were able to discriminate standards from deviants in the posttest (> 60% correct responses). The remaining 26 participants were classified as nonlearners. Learners and nonlearners did not differ significantly with respect to age and sex (learners: mean age = 24.3 y, SD = 2.9, 50% females; nonlearners: mean age = 24.9 y, SD = 2.4, 54% females).

**Stimuli.** The syllables were spoken separately by a trained female speaker. To have naturally sounding syllables, we recorded many exemplars and selected examples which were similar in pitch and length. Each syllable was stored in a separate file of 250-ms duration.

We used a total of 818 stimuli, with 658 standard stimuli and 80 tokens of each deviant type (for fundamental frequencies of all syllables, see Table S2 and Audio S1).

**Procedure.** The procedure was an oddball-paradigm with standard stimuli occurring in ~80% of the trials and rule and pitch deviants in ~10% of the trials. Syllables were separated from their neighbors within a trisyllabic sequence by a gap of 50 ms and there was a gap of 700 ms between sequences. Stimuli were presented in pseudorandomized order, which ensured that each A...B rule occurred with equal frequency in a given series of standards, and that the identical A...B rule was not repeated more often than three times. Between two deviants, sequences consisting of two, four, six, or eight standard stimuli occurred. When the sequence of standard stimuli was  $\leq 4$ , we ensured that different types of deviants occurred before and afterward to provide enough exemplars to re-establish the standard A...B rules.

After the ERP experiment, adult participants were asked to perform a familiarity judgment for 40 standards and 40 rule deviants. During each trial, a single AXB triplet was presented and participants had to decide whether the sequence sounded like a "frequently occurring sequence" or an "infrequently occurring sequence."

**EEG Recording.** The continuous EEG was recorded from Ag/AgCl electrodes (27 in infants, 61 in adults), fixed in an elastic cap placed on the participants' head (infants: EASYCAP; adults: ElectroCap International), with a sampling rate of 500 Hz. Adult participants and caregivers holding the infant sat in a sound-proof booth during stimulus presentation. Stimuli were played via two loudspeakers at a comfortable sound level constant across subjects. Online reference was Cz in infants and the left mastoid in adults. Offline, the EEG data were rereferenced to linked mastoids, band-pass-filtered between 0.3 and 20 Hz (−3 dB, cutoff frequencies of 0.38 Hz and 19.92 Hz), epoched, and averaged from −100 to 800 ms relative to the onset of the final syllable of the trisyllabic sequences, with a baseline of −100 to 0 ms (EEP 3.2., ANT-software). In infants, artifacts were rejected after automatic (trials exceeding a SD of 70  $\mu$ V in a sliding window of 500 ms) and manual inspection. Only infants with more than 20 trials per condition were included in the analysis. In adults, the software EEGLAB 6.01 (37) was used for artifact correction.

After manual rejection of large artifacts, the continuous data were entered into an independent component analysis. The resulting components were used to reject ocular artifacts.

**Data Analysis. Experiment 1.** For the infants, we first applied a categorization criterion based on spatial, temporal, and amplitude parameters of the individual ERP response. We chose eight electrodes from frontocentral locations (C3, C4, CZ, F3, F4, FC5, FC6, FZ) as indicator electrodes for the presence of a negativity. An infant was assigned to the negativity group whenever there was an amplitude difference of  $< -0.5 \mu\text{V}$  between deviants and standards for at least four electrodes of the indicator set, and in at least three subsequent 20-ms-long TWs between 60 and 260 ms after stimulus onset.

For the statistical analysis, 15 representative electrodes were assigned to the different levels of the factors LATERALITY (left: F7, FC5, CP5; left medial: F3, C3, P3; medial: FZ, CZ, PZ; right medial: F4, C4, P4; right: F8, FC6, CP6) and REGION (anterior: F7, F3, FZ, F4, F8; central: FC5, C3, CZ, C4, FC6; posterior: CP5, P3, PZ, P4, CP6). Mean amplitudes were entered into five-way ANOVAs with the between-subjects factors SEX (female vs. male) and MMR (neg MMR vs. pos MMR for the pitch condition), and the within-subject factors LATERALITY (five levels from left to right) and REGION (three levels from anterior to posterior) and RULE (rule deviant vs. standard) or PITCH (pitch deviant vs. standard), respectively. Greenhouse–Geisser-corrected *P* values are reported for all experiments whenever degrees of freedom are  $>1$ . To determine if there were significant effects for the pitch condition and the rule condition, we first calculated 20-ms running TW analyses. Whenever  $P < 0.05$  for an

effect including RULE or PITCH was confirmed in four or more consecutive TWs, subsequent ANOVAs were calculated across the whole TW. To estimate the signal-to-noise ratio across the experimental groups, we conducted ANOVAs on the number of averaged trials in each experimental condition. There were no significant effects (all *P*s  $> 0.5$ ). On average, there were 219 (SD 67) trials for standards, 44 (SD 10) trials for the pitch condition and 42 (SD 10) trials for the rule condition.

**Experiments 2 and 3.** The adult ERP data were evaluated with the same statistical model as the infant data, but more electrodes were included within each region of interest. Thirty representative electrodes were assigned to the different levels of the factors LATERALITY (left: F5, FC5, C5, CP5, P5, PO7; left medial: F3, FC3, C3, CP3, P3, PO3; medial: Fz, FCz, Cz, CPz, Pz, POz; right medial: F4, FC4, C4, CP4, P4, PO4; right: F6, FC6, C6, CP6, P6, PO8) and REGION (anterior: F5, FC5, F3, FC3, Fz, FCz, F4, FC4, F6, FC6; central: C5, CP5, C3, CP3, Cz, CPz, C4, CP4, C6, CP6; posterior: P5, PO7, P3, PO3, Pz, POz, P4, PO4, P6, PO8). In Exp. 3, the additional between-subject factor GROUP (learners vs. non-learners) was introduced. For Exps. 2 and 3, we conducted additional regression analyses (*SI Regression Analysis*).

**ACKNOWLEDGMENTS.** We thank all the participating infants and parents, Anna S. Hasting for support in designing the study, the baby laboratory team and Galina Surova for data acquisition and evaluation, and Uta Frith and Hannes Rakoczy for helpful comments on previous versions of the manuscript. The project was funded by Grants MU 3112/1-1 and MU 3112/1-2 from the Deutsche Forschungsgemeinschaft.

- Tsao FM, Liu HM, Kuhl PK (2004) Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Dev* 75:1067–1084.
- Leppänen PHT, et al. (2010) Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex* 46:1362–1376.
- Jusczyk PW (1997) *The Discovery of Spoken Language* (MIT Press, Cambridge, MA).
- Wible B, Nicol T, Kraus N (2005) Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain* 128:417–423.
- Jakoby H, Goldstein A, Faust M (2011) Electrophysiological correlates of speech perception mechanisms and individual differences in second language attainment. *Psychophysiology* 48:1517–1531.
- Díaz B, Baus C, Escera C, Costa A, Sebastián-Gallés N (2008) Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proc Natl Acad Sci USA* 105:16083–16088.
- Saffran JR, Aslin RN, Newport EL (1996) Statistical learning by 8-month-old infants. *Science* 274:1926–1928.
- Teinonen T, Fellman V, Näätänen R, Alku P, Huotilainen M (2009) Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neurosci* 10:21.
- Gervain J, Macagno F, Cogoi S, Peña M, Mehler J (2008) The neonate brain detects speech structure. *Proc Natl Acad Sci USA* 105:14222–14227.
- Friederici AD, Mueller JL, Oberecker R (2011) Precursors to natural grammar learning: Preliminary evidence from 4-month-old infants. *PLoS ONE* 6:e17920.
- Thiessen ED, Hill EA, Saffran JR (2005) Infant-directed speech facilitates word segmentation. *Infancy* 7:53–71.
- Langus A, Marchetto E, Bion RAH, Nespor M (2012) Can prosody be used to discover hierarchical structure in continuous speech? *J Mem Lang* 66:285–306.
- Mueller JL, Bahlmann J, Friederici AD (2010) Learnability of embedded syntactic structures depends on prosodic cues. *Cogn Sci* 34:338–349.
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270:303–304.
- Trainor LJ, Desjardins RN (2002) Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychon Bull Rev* 9:335–340.
- Gómez RL (2002) Variability and detection of invariant structure. *Psychol Sci* 13:431–436.
- Gómez RL, Maye J (2005) The developmental trajectory of nonadjacent dependency learning. *Infancy* 7:183–206.
- Lany J, Gómez RL (2008) Twelve month-old infants benefit from prior experience in statistical learning. *Psychol Sci* 19:1247–1252.
- Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin Neurophysiol* 118:2544–2590.
- Paavilainen P, Simola J, Jaramillo M, Näätänen R, Winkler I (2001) Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). *Psychophysiology* 38:359–365.
- Cheour M, et al. (2002) Speech sounds learned by sleeping newborns. *Nature* 415:599–600.
- Ceponiene R, et al. (2002) Event-related potential features indexing central auditory discrimination by newborns. *Brain Res Cogn Brain Res* 13:101–113.
- He C, Hotson L, Trainor LJ (2007) Mismatch responses to pitch changes in early infancy. *J Cogn Neurosci* 19:878–892.
- He C, Hotson L, Trainor LJ (2009) Maturation of cortical mismatch responses to occasional pitch change in early infancy: Effects of presentation rate and magnitude of change. *Neuropsychologia* 47:218–229.
- Kushnerenko E, Ceponiene R, Balan P, Fellman V, Näätänen R (2002) Maturation of the auditory change detection response in infants: A longitudinal ERP study. *Neuroreport* 13:1843–1848.
- Galsworthy MJ, Dionne G, Dale PS, Plomin R (2000) Sex differences in early verbal and non-verbal cognitive development. *Dev Sci* 3:206–215.
- Vouloumanos A, Werker JF (2004) Tuned to the signal: The privileged status of speech for young infants. *Dev Sci* 7:270–276.
- De Diego Balaguer R, Toro JM, Rodriguez-Fornells A, Bachoud-Lévi A-C (2007) Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS ONE* 2:e1175.
- Friederici AD, Friedrich M, Weber C (2002) Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *Neuroreport* 13:1251–1254.
- Xu L, Thompson CS, Pfingst BE (2005) Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am* 117:3255–3267.
- Bendixen A, Schröger E, Winkler I (2009) I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. *J Neurosci* 29:8447–8451.
- Friederici AD, et al. (2008) Sex hormone testosterone affects language organization in the infant brain. *Neuroreport* 19:283–286.
- Eimer M, Goschke T, Schlaghecken F, Stürmer B (1996) Explicit and implicit learning of event sequences: Evidence from event-related brain potentials. *J Exp Psychol Learn Mem Cogn* 22:970–987.
- Kramer JH, Delis DC, Kaplan E, O'Donnell L, Prifitera A (1997) Developmental sex differences in verbal learning. *Neuropsychology* 11:577–584.
- Ferdinand NK, Mecklinger A, Kray J (2008) Error and deviance processing in implicit and explicit sequence learning. *J Cogn Neurosci* 20:629–642.
- Birdsong D, Molis M (2001) On the evidence for maturational constraints in second-language acquisition. *J Mem Lang* 44:235–249.
- Delorme A, Makeig S (2004) EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21.

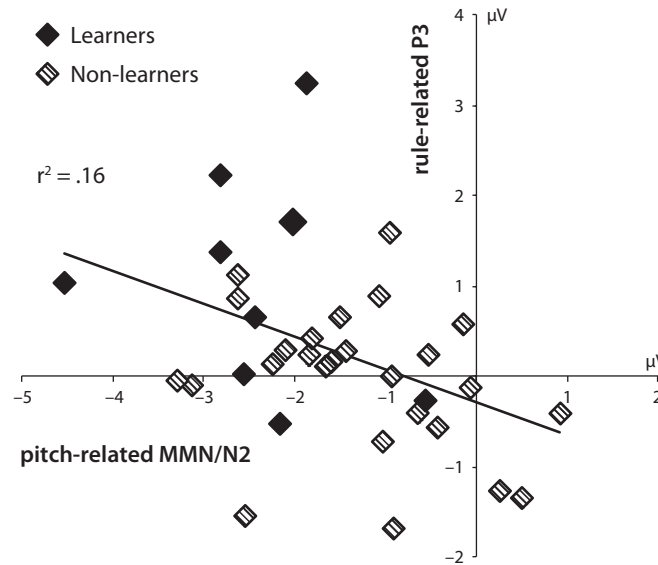
# Supporting Information

Mueller et al. 10.1073/pnas.1204319109

## SI Regression Analysis

To make the statistical approach to adults' and infants' data more comparable, we additionally conducted regression analyses for Exps. 2 and 3 with mean amplitudes of the pitch condition (pitch deviants – standards) as predictors and mean amplitudes of the rule condition (rule deviants – standards) as criterion. For Exp. 2, this was done for the two time windows (TWs) in which we found significant effects for the pitch condition. For Exp. 3, this

was done for the observed mismatch negativity (MMN)/N2 and P3 effects in the pitch and the rule conditions. In Exp. 2 there was no significant result. In Exp. 3 the pitch MMN/N2 significantly predicted the amplitude of the P3 ( $r^2 = 0.16$ ,  $P < 0.01$ ) (Fig. S1) but not the amplitude of the N2. Thus, the rule-related P3 was predicted by the pitch-related MMN/N2 across both the learner and the nonlearner group when participants were engaged in an active task.



**Fig. S1.** Mean amplitude of the MMN/N2 for pitch processing predicts amplitude of the P3 in the rule condition in Exp. 3. This indicates a linear relationship between aspects of pitch perception and rule detection across all adult participants in an active target detection task.

**Table S1. Additional physiological participant information**

Participant group	Age (d)	Gestational age (wk)	Birth weight (g)
Girls: Neg pitch MMR ( $n = 15$ )	91 (SD 10)	39.9 (SD 1.1)	3,372 (SD 286)
Boys: Neg pitch MMR ( $n = 18$ )	93 (SD 7)	39.4 (SD 0.9)	3,431 (SD 262)
Girls: Pos pitch MMR ( $n = 17$ )	92 (SD 9)	39.8 (SD 1.3)	3,518 (SD 503)
Boys: Pos pitch MMR ( $n = 15$ )	89 (SD 9)	40.1 (SD 1.3)	3,634 (SD 262)

For all three dependent variables, two-factorial ANOVAs with the between-subject factors SEX (male vs. female) and mismatch response (MMR) (neg MMR group vs. pos MMR group) were calculated. There were no significant effects (all  $P > 0.5$ ).

**Table S2. Mean intensity (dB) and pitch maximum (Hz) of auditory stimuli**

Syllable	Intensity	Pitch
<b>bu</b>	<b>86</b>	<b>277</b>
<b>fi</b>	<b>87</b>	<b>282</b>
gi	88	275
gu	88	291
ka	86	273
ke	86	275
ki	87	279
ko	86	278
ku	86	272
<b>le</b>	<b>85</b>	<b>269</b>
ma	84	269
me	86	264
mi	88	273
mu	88	274
no	87	273
ri	85	267
ro	86	276
sa	85	275
so	86	279
su	87	278
<b>to</b>	<b>87</b>	<b>277</b>
wa	85	269
we	86	263
xe	85	285
<b>*bu</b>	<b>86</b>	<b>304</b>
<b>*to</b>	<b>87</b>	<b>305</b>

A and B syllables from the AXB sequences are written in boldface. The syllables with manipulated pitch are marked with an asterisk.

**Audio S1.** Example stimuli corresponding to the syllables depicted in Fig. 1.

[Audio S1](#)