

# Non-adjacent auditory sequence learning across development and primate species

Jutta L. Mueller<sup>1</sup>, Alice Milne<sup>2,3,4</sup> & Claudia Männel<sup>5,6,7</sup>

*1 Institute of Cognitive Science, University of Osnabrück, Wachsbleiche 27, 49090 Osnabrück Germany*

*2 Ear Institute, University College London, 332 Grays Inn Road, London, WC1X 8EE, United Kingdom*

*3 Institute of Neuroscience, Henry Wellcome Building, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom*

*4 Centre for Behaviour and Evolution, Henry Wellcome Building, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, United Kingdom*

*5 Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany*

*6 Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany*

*7 Clinic for Cognitive Neurology, Medical Faculty, University of Leipzig, Liebigstr. 16, 04103 Leipzig, Germany*

## **Corresponding author:**

Jutta L. Mueller  
University of Osnabrück  
Institute of Cognitive Science  
Wachsbleiche 27  
49090 Osnabrück  
Germany  
Email: [jutta.mueller@uni-osnabrueck.de](mailto:jutta.mueller@uni-osnabrueck.de)  
Tel.: +49 541 969 3369

***Research Highlights:***

- Non-adjacent dependencies (NADs) are central building blocks of human language
- Electrophysiology shows infants' early learning of NADs in auditory sequences
- Electrophysiology shows similar NAD learning in non-human primates and human infants
- Auditory sequencing seems to work as a scaffold for the evolution of human syntax
- Infants often outperform adults in automatic rule extraction from auditory sequences

### ***Abstract***

The ability to process structured sequences of sounds lies at the basis of human language processing. Language is characterized by a high level of structural complexity including non-adjacent dependencies where the relationships between elements can span multiple intervening elements. Understanding how such structures can be learned is of particular interest for understanding how language develops and how it evolved. In recent years advances have been made in identifying the constraints and conditions under which non-adjacent dependencies can be learned across development and different species. Yet, in humans, online and offline methods suggest markedly different developmental timelines. Notably, human infants seem to outperform adults in online learning during mere exposure. Comparative studies with non-human primates reveal the auditory sequence learning capacities and limitations of our close phylogenetic relatives. Initial findings suggest that primates show similar learning abilities to human infants revealing an evolutionary preserved computational ability to automatically extract non-adjacent relations from auditory sequences. The pattern found across ontogeny and phylogeny is consistent with the idea that language evolution was in part enabled by powerful auditory sequencing abilities. These abilities were potentially boosted in humans by improvements in higher-order cognitive abilities that allowed us to link sequential patterns to abstract semantic and syntactic representations.

When we speak, we do something that is very widespread among animals. We produce a spectrally and temporally complex sequence of vocalizations. While human language additionally involves combinatorial structure and semantic meaning, there are ongoing debates concerning the degree of syntactic structure and semantics present in other animals [cf. 1–4, for a controversial discussion]. Independent of whether certain precursors of language exist in animals, the boundless nature of language [5,6] may be one reason why it became such an incredibly useful tool, enabling human cognition and communication [7]. Humans are not born with a fully developed language system, but display marked developmental changes and transitions across the relevant cognitive and neural processes [8]. Supporting human linguistic ability is the auditory perception system, that is not specific to language per se, but contributes to how it is perceived (e.g. through hearing range, grouping biases, and temporal resolution). Note that this is not a claim about how language evolved with respect to potential gestural vs. vocal origins [e.g. 9,10]. It simply describes the sensory basis of the signal which serves as the carrier for the most widespread and default form of language - that is, spoken language. These basic auditory processes likely predate the emergence of language phylogenetically [11], and should be evident in species that are evolutionarily related to humans [12]. Therefore, we suggest that a full understanding of language and its evolution requires a developmental-comparative perspective, as has been called for in other fields of psychology [13,14].

The present review focuses on the processing of non-adjacent dependencies (NADs) as a specific domain of auditory sequence processing. NADs are of crucial importance to human language, which cannot be fully described, syntactically, by fixed patterns or combinations of adjacent speech sounds. Syntactic relations frequently span a considerable distance. For example, the sentence “Daddy who is the best storyteller in the world always tells me stories about knights” contains a long-distance dependency between the subject, Daddy, and the third-person suffix –s on the verb tell (i.e. tells). The linking of non-adjacent linguistic units across a distance presupposes the availability of sufficient memory and processing capacities, on the part of the listener, both for learning and processing. Note that at the level of linguistic description, NADs may become adjacent dependencies. This is because different sentence parts are analysed in larger, hierarchically structured units, such as noun and verb phrases. Moreover, the same NAD can, in principle, be analysed at different levels of abstraction. In the above example, an infant might notice that the word Daddy always leads to a suffix –s somewhat later, while a proficient language user has syntactic knowledge about the categorical relationship between subjects and third person agreement on the verb. Despite these caveats on non-adjacency and levels of abstraction, language learners undoubtedly process NADs at the input-level first, before mapping them onto more abstract syntactic levels of representation. For this review we restrict the discussion to NADs in sequences of speech sounds in the absence of syntactic function and semantic meaning. How the capacity to detect and learn NADs in the sequential speech stream develops in

Deleted: r

early infancy has been a topic of intense investigation lately. In the present report we summarize and interpret recent results on the processing of NADs across human development as well as in non-human primates. Our aim is to shed light on whether, and how this ability potentially served as a precursor of the uniquely human ability to learn language.

### **Human developmental studies**

#### *Behavioural*

Developmental studies have often used the head-turn preference procedure to study NAD learning, using either the child's native language or an artificial grammar. These paradigms measure infants' natural inclination to orientate differently towards familiar and novel types of stimuli [15]. Behavioural sensitivity to NADs between native language morphemes, for example in auxiliary-verb congruence of gender agreement, is evident after the first birthday [16–19]. For artificial grammars, behavioural evidence shows children's sensitivity to NADs emerging around 12 months of age and stabilising around 18 months [20–23]. There is some evidence of NAD sensitivity before that age, but only in the limited domain of repetition detection, which does not necessarily involve the learning of specific relations between non-adjacent elements [24]. Linking artificial grammar learning to language acquisition, there is now evidence that the ability to learn NADs is impaired in developmental language disorders [23]. In sum, behavioural studies indicate that sensitivity to NADs, between different speech segments or morphemes, emerges around one year of age.

#### *Electrophysiological*

In contrast to behavioural measures, electrophysiological indices, measured during learning or processing of NADs, do not depend on overt responses and thus allow for a more direct comparison of infant and adult learning processes. In a series of experiments using event-related potentials (ERPs), we tested whether monolingual, German infants and adults were able to learn NADs from a natural, non-native language [25–28]. Participants were exposed to correct Italian sentences during a learning phase and subsequently tested on novel stimuli that either followed or violated the NAD (cf. Figure 1 A, B). Most interestingly, infants of only 4 months of age proved sensitive to the difference between correct and incorrect NADs, indicating that they had extracted the grammatical regularity in the learning phases (cf. Figure 1 C) [25]. When adults were tested in the same way as the infants, that is during passive listening in the absence of any task, there was no indication of learning (cf. Figure 1 D) [27].

--- please insert Figure 1 here ---

In a different study, Mueller et al. [29] adapted an oddball design to test whether and how auditory perception and the learning of NADs are linked in infancy and adulthood (cf. Figure 2). Passive listening oddball paradigms are designed such that frequently occurring standard stimuli (e.g., 80%) are interspersed with rare stimuli (e.g., 20%), that deviate on one stimulus dimension from the standard stimuli [cf. 30, for review]. The subtraction of the ERP waveforms related to the standards from those related to the deviants provides an attention-independent marker of stimulus discrimination, the Mismatch Negativity (MMN), which, in its immature form in young infants, is often reversed in polarity [31]. The oddball paradigm used in Mueller et al. [29] featured syllable-based NADs between two elements (A and B), in AXB sequences, as the standard stimuli (cf. Figure 2 A). Two different deviant conditions were used to probe the relationship between NAD learning and basic auditory perception. More specifically, MMNs could arise from violations of the dependency between A and B or from pitch increases placed on the B syllable. ERP analysis revealed that 3-month-olds with a mature MMN, as measured in the pitch condition, were also sensitive to the NADs. However, infants who showed the immature MMN were not sensitive to NADs. That is, detection of NAD violation required a mature MMN in response to basic auditory oddball stimuli (cf. Figure 2 C). In contrast, adults did not show any NAD learning effect during passive listening, despite responding to the change in pitch (cf. Figure 2 D).

Importantly, in all of these NAD learning experiments, adults who were given active judgment or detection tasks, showed some learning success reflected in a biphasic ERP pattern of a negativity and a subsequent positivity (cf. Figure 1 E) [26,28,29]. Interestingly, as was the case for young infants, successful adult learners showed an enhanced pitch-related ERP response compared to unsuccessful learners [29]. Thus, the ability to learn NADs seems to be closely linked to auditory processes across development, over and above the general developmental decline of sensitivity in passive learning situations.

This series of ERP studies led to the following conclusions. Infants are sensitive to NADs even at 3 months of age. Adults have lost this sensitivity to some degree, in passive learning settings, but are still able to learn when given a judgment task. Based on this developmental pattern we hypothesized that frontal cortex-based functions of cognitive control that mature around the age of four years, may cause a fundamental change in learning mechanisms, from automatic associative learning in infancy, towards controlled, attention-guided learning in adulthood [32]. In support of this theory, an ERP study on the learning of NADs in preschool children revealed that the age of 4 years may be a developmental turning point [33]. Furthermore, in adults it is possible to influence the brain regions that drive cognitive control. For example, application of transcranial direct current stimulation (TDCS)

to dorsolateral prefrontal cortex during NAD learning resulted in a change of the electrophysiological responses, although behavioural effects were not observed [27].

Deleted: s

A related set of studies used frequency tagging of EEG-oscillations to test the segmentation of words containing NADs in infants and adults [34,35]. Adults exposed to trisyllabic sequences containing NADs displayed both, a power decrease for the frequency reflecting the unisyllabic presentation rate, and a power increase for the frequency corresponding to the duration of the trisyllabic units compared to randomly concatenated syllables. Importantly, the latter effect was correlated with word recall [34]. Infants exposed to NADs in a similar experimental design showed increased phase-locking for the word-length frequency [35]. These results show that the human brain entrains to the frequency representing the major units that contain NADs during learning. Time-frequency analysis thus provides additional, interesting measures for online learning processes.

Taken together, electrophysiological evidence suggests sensitivity to NADs develops much earlier (from 3 months) than behavioural, attention-dependent measures indicate. While adults can learn NADs in principle, the available data suggest that the learning mechanisms employed in infants and adults differ. Infants outperform adults in passive listening situations. Thus, development can be characterized by a shift from automatic sampling of frequent input patterns to controlled, task-dependent stimulus selection. The developmentally early ability to identify NADs suggests that the processing of long-distance relationships is an inherent aspect of basic human cognition, which may serve as a scaffolding for children's first steps into language. These developmental changes provide a unique opportunity to contrast the abilities of species that represent an evolutionarily earlier cognitive stage, with infants, who provide an early ontogenetic perspective. Therefore, we should consider whether a similar capacity is also present in our primate relatives and other species.

## Primate studies

### *Behavioural*

Recently, an increasing number of behavioural studies have tested different non-human primate species on their sensitivity to a range of NADs [36–40]. Chimpanzees, our closest relatives, can correctly identify the relative order of non-adjacent number positions (presented visually), despite prior training with only adjacent number sequences [39]. They can also detect arbitrary or similarity-based relationships between non-adjacent shapes [40]. In contrast to the visual domain, there is limited direct evidence for NAD processing in the auditory domain in non-human primates. Endress and colleagues reported chimpanzee's sensitivity to items at edge positions of 6-unit sequences consisting of 3 different species-specific call types, but there was no indication that the animals established relationships between distant items [41]. Furthermore, macaques showed little

sensitivity to NADs in human speech stimuli, computer-generated sounds or visual shapes [38,42]. Interestingly, squirrel monkeys, more distantly related primates, reportedly learned the correspondence between the category of the first and last elements of an acoustic sequence, across a variable distance [36]. However, this finding has been questioned given methodological weaknesses [43] and the lack of experimental conditions that show the generalization of NADs beyond positional rules between instances of a perceptual category. The only primate species to provide behavioural evidence for NAD sensitivity, is the tamarin monkey, a species whose phylogenetic lineage diverged from the human lineage a long time ago. Newport and colleagues reported that tamarin monkeys can successfully learn NADs between elements of human speech, specifically, dependencies between two syllables and two vowels, but not between consonants [37]. In contrast, human learners who were given the task of finding word boundaries showed sensitivity to consonant and vowel-based dependencies, but not syllables [44]. This pivotal study in tamarin monkeys provides the first evidence that the ability to compute NADs might be present in the primate lineage, yet governed by different constraints compared to humans.

#### *Electrophysiological*

Linking the human electrophysiological studies discussed above to the comparative work in non-human primates, Milne et al. [45] conducted an ERP study on NAD learning in macaque monkeys. The experiment was closely modelled on the aforementioned human oddball study [29], and tested macaques' sensitivity to pitch variations and to NADs between specific syllables and their non-adjacent counterpart in an AXB grammar (cf. Figure 2 A). Previous studies have shown auditory MMNs in primates but only in response to neutral, non-human sounds [46]. Interestingly, in Milne et al. [45], macaques showed an MMN response to pitch variations in human speech and also to NAD violations. Their responses occurred in a similar time window as had been previously found in human infants [29].

--- please insert Figure 2 here ----

Importantly, human adults did not show any passive learning with these kinds of stimuli and when they did learn, when given an explicit task, their ERP responses occurred later than both the infants and the monkeys. Thus, during mere exposure to NADs, macaque monkeys show a sensitivity that is more similar to human infants than human adults. In sum, the ability to extract NADs from passive input sampling is present early in phylogeny as well as early in ontogeny, undergoing modification across human development. These findings support theoretical accounts of the emergence of language, where the ability to extract sequentially presented relationships is considered a scaffold



for the evolution of human syntax [11, 47]. In humans, this ability undergoes major developmental transformations, which allow for the mapping of syntactic and semantic information.

## Conclusion

Recent electrophysiological studies on NAD learning in human and non-human primates go beyond behavioural evidence and suggest that the ability to extract NADs from auditory sequences might be present early in phylogeny, as well as early in ontogeny. Moreover, current evidence points to human-specific sensitivities as well as the existence of marked developmental changes, yet to be found in our primate relatives. A comparative-developmental perspective, as suggested here, can help to specify what exactly is shared or divergent between species at behavioural and neural levels.

What exactly triggers the human developmental changes and results in differences across primate species is an open research question. The development of higher cognitive functions related to cognitive control are plausible candidates. Neurodevelopmental models of language, put forth by Skeide and Friederici [8] for example, stress the role of the dorsal fibre pathway for the development of language and thus, provide a framework for integrating the observed developmental and cross-species differences. In order to explain the human capacity for language, as a whole, we need to broaden the focus and explain how general sequence learning ability leads to the mapping of syntactic function and semantic meaning, two core ingredients of language. While the current data suggest that some sequence learning abilities are hampered by cognitive control, such as automatic NAD learning during mere listening, the opposite seems true for advanced language learning, because function and meaning mapping rely on the mental workspace provided by cognitive control. Nonetheless, the relative decline of automatic learning processes after childhood might provide an explanation for the long-standing concept of a critical or sensitive-period hypothesis for language acquisition as formulated, for example, by Lenneberg [48], which states that language learning after a certain developmental stage is impaired and qualitatively changed. Future comparative-developmental work will help in substantiating, refining or discarding these ideas, which are based on the intriguing parallels between electrophysiological correlates of auditory sequence processing in human infants and adult macaques.

**Box 1: Prospects and pitfalls of comparative EEG studies on NADs****Prospects**

NADs have yet to be fully explored using comparative EEG. Previous work has considered ERP patterns associated with sequence processing across primates [45,49]. Yet, recent invasive neural recordings suggest that adjacent sequence learning modulates oscillatory coupling in both humans and monkeys [50]. This has yet to be fully investigated using surface EEG recordings and applied to NADs. In macaques it would also be possible to further investigate the neural underpinnings of the MMN elicited in response to NAD violations via direct neural recordings from frontal and temporal cortices.

**Pitfalls**

It is important to consider the limitations of comparative studies, particularly with regard to null results. Learning of NADs in non-human primates has been found only for a limited number of dependencies in human speech [37]. The pattern may look different for primates' own calls or non-biological sounds. Human and animal studies will often have unavoidable methodological variation, for example, in terms of sample size, stimulation duration or testing environment. Technical advances, for example, the development of fully portable head-mounted EEG systems that are less sensitive to movement artifacts, will open up new avenues to test human infants and animals under more comparable conditions.

**Box 2: Open questions for future research**

- What is the relevance of early developmental stages of sequence learning for language acquisition?
- What triggers the developmental pattern in human sequence learning?
- What is the upper limit of structural complexity in non-human primate and human infant NAD learning?
- What are the perceptual and computational constraints for the elements coding NADs across species?

## References

1. Townsend SW, Koski SE, Byrne RW, Slocombe KE, Bickel B, Boeckle M, Braga Goncalves I, Burkart JM, Flower T, Gaunet F, et al.: **Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals.** *Biol Rev* 2017, **92**:1427–1433.
2. Fischer J: **Primate vocal production and the riddle of language evolution.** *Psychon Bull Rev* 2017, **24**:72–78.
3. Abe K, Watanabe D: **Songbirds possess the spontaneous ability to discriminate syntactic rules.** *Nat Neurosci* 2011, **14**:1067-U173.
4. Beckers GJL, Bolhuis JJ, Okanoya K, Berwick RC: **Birdsong neurolinguistics: songbird context-free grammar claim is premature.** *Neuroreport* 2012, **23**:139–45.
5. von Humboldt W: *On Language: On the Diversity of Human Language Construction and Its Influence on the Mental Development of the Human Species.* Cambridge University Press (Originally published in 1836); 1836.
6. Chomsky N: *Syntactic structures.* The Hague: Mouton; 1957.
7. Vygotski LS: *Thought and language.* MIT Press; 1986.
8. Skeide MA, Friederici AD: **The ontogeny of the cortical language network.** *Nat Rev Neurosci* 2016, **17**:323–332.
9. Arbib MA: **From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics.** *Behav Brain Sci* 2005, **28**.
10. Studdert-Kennedy M: **Evolutionary implications of the particulate principle: imitation and the dissociation of phonetic form from semantic function.** In *The Evolutionary Emergence of Language.* Edited by Knight C, Studdert-Kennedy M, Hurford JR. Cambridge University Press; 2000:161–176.
11. Christiansen MH, Chater N: **Language as shaped by the brain.** *Behav Brain Sci* 2008, **31**:489–509.
12. Hauser MD, Chomsky N, Fitch WT: **The faculty of language: What is it, who has it, and how did it evolve?** *Science* 2002, **298**:1569–1579.
13. Nielsen M, Haun D: **Why developmental psychology is incomplete without comparative and cross-cultural perspectives.** *Philos Trans R Soc B Biol Sci* 2016, **371**:20150071.

14. Beck SR: **Interaction between comparative psychology and cognitive development.** *Curr Opin Behav Sci* 2017, **16**:138–141.
15. Kemler Nelson DG, Jusczyk PW, Mandel DR, Myers J, Turk A, Gerken L: **The head-turn preference procedure for testing auditory perception.** *Infant Behav Dev* 1995, **18**:111–116.
16. Hoehle B, Schmitz M, Santelmann LM, Weissenborn J: **The Recognition of Discontinuous Verbal Dependencies by German 19-Month-Olds: Evidence for Lexical and Structural Influences on Children’s Early Processing Capacities.** *Lang Learn Dev* 2006, **2**:277–300.
17. Santelmann LM, Jusczyk PW: **Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space.** *Cognition* 1998, **69**:105–134.
- \*\*18. Culbertson J, Koulaguina E, Gonzalez-Gomez N, Legendre G, Nazzi T: **Developing knowledge of nonadjacent dependencies.** *Dev Psychol* 2016, **52**:2174–2183.

Infants of 6 different age groups between 14 and 24 months were exposed to French sentences containing NADs in the form of subject-verb agreements. Infants’ looking time differences to grammatical and ungrammatical (with agreement violations) stimuli revealed a developmental progression from infants’ familiarity preference of grammatical to a novelty preference of ungrammatical sentences. This finding can be viewed as increasing robustness of infants’ representations based on continued language exposure. Most interestingly, infants’ looking behaviour indicated that this developmental progression occurs twice in the tested age range. This finding is particularly appealing, as it might indicate a change in the underlying representations between 18 and 21 months, from surface-based phonological to feature-based morpho-syntactic patterns. Thus, fine-grained testing of narrow age ranges can uncover complex developmental trajectories.

19. van Heugten M, Shi R: **Infants’ sensitivity to non-adjacent dependencies across phonological phrase boundaries.** *J Acoust Soc Am* 2010, **128**:EL223-EL228.
20. Gómez RL, Maye J: **The developmental trajectory of nonadjacent dependency learning.** *Infancy* 2005, **7**:183–206.
21. Lany J, Gómez RL: **Twelve-Month-Old Infants Benefit From Prior Experience in Statistical Learning.** *Psychol Sci* 2008, **19**:1247–1252.
22. Gómez RL: **Variability and detection of invariant structure.** *Psychol Sci* 2002, **13**:431–436.
23. Kerkhoff A, De Bree E, de Klerk M, Wijnen F: **Non-adjacent dependency learning in infants at familial risk of dyslexia.** *J Child Lang* 2013, **40**:11–28.

24. Gervain J, Werker JF: **Learning non-adjacent regularities at age 0 ; 7. *J Child Lang* 2013, 40:860–72.**
25. Friederici AD, Mueller JL, Oberecker R: **Precursors to Natural Grammar Learning: Preliminary Evidence from 4-Month-Old Infants. *PLoS One* 2011, 6:e17920.**
26. Mueller JL, Oberecker R, Friederici AD: **Syntactic learning by mere exposure An ERP study in adult learners. *BMC Neurosci* 2009, 10:89.**
27. Friederici AD, Mueller JL, Sehm B, Ragert P: **Language Learning without Control: The Role of the PFC. *J Cogn Neurosci* 2013, 25:814–821.**
28. Citron FMM, Oberecker R, Friederici AD, Mueller JL: **Mass counts: ERP correlates of non-adjacent dependency learning under different exposure conditions. *Neurosci Lett* 2011, 487:282–6.**
- \*\*29. Mueller JL, Friederici AD, Männel C: **Auditory perception at the root of language learning. *Proc Natl Acad Sci USA* 2012, 109:15953–8.**

Three-month-old infants track non-adjacent dependencies in an artificial grammar learning paradigm. Adults tested under the same passive exposure do not show learning, except if given a grammaticality judgement task. There seems to be a developmental shift in learning mechanisms from infants' automatic sampling of structured input to adults' more effortful, attention-guided input decoding. Crucially, electrophysiological measures predate children's capacity to track non-adjacent dependencies in the auditory domain to early infancy, in contrast to behavioural measures that evidence this ability only starting in the second year of life. See [25], for similar evidence.

30. Näätänen R, Paavilainen P, Rinne T, Alho K: **The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin Neurophysiol* 2007, 118:2544–2590.**
31. He C, Hotson L, Trainor LJ: **Development of infant mismatch responses to auditory pattern changes between 2 and 4 months old. *Eur J Neurosci* 2009, 29:861–7.**
32. Thompson-Schill SL, Ramscar M, Chrysikou EG: **Cognition Without Control: When a Little Frontal Lobe Goes a Long Way. *Curr Dir Psychol Sci* 2009, 18:259–263.**
- \*33. Mueller JL, Friederici AD, Männel C: **Developmental changes in automatic rule learning mechanisms across early childhood. in revision.**

Electrophysiological markers of NAD learning show a developmental change during preschool age that leads away from the infant-specific, automatic tracking of distant input relations. Specifically, at 2 years of age, children can still learn NADs in an artificial grammar from passive exposure, while

children at 4 years have largely lost this automatic ability. This developmental change is likely caused by prefrontal cortex maturation during this age, with increased cognitive control shifting learning capacities from automatic input sampling to attention-guided decoding strategies.

34. Buiatti M, Peña M, Dehaene-Lambertz G: **Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses.** *Neuroimage* 2009, **44**:509–519.
- \*35. Kabdebon C, Pena M, Buiatti M, Dehaene-Lambertz G: **Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants.** *Brain Lang* 2015, **148**:25–36.

Infants at 8 months of age were exposed to three-syllabic test words that were consistent (e.g., ku-na-bi) with a non-adjacent rule introduced in a familiarization speech stream (i.e., ku...bi) as compared to test words that were inconsistent with this rule (e.g., fi-bi-na). The analyses of event-related brain potentials and EEG phase-locking values indicate infants' differential processing of consistent and inconsistent test words. The results indirectly indicate infants' potential extraction of NADs. Note that the methodological approach to combine ERP and time-frequency analysis is promising. Further, the fact that there were no processing differences between full-term and pre-term-infants confirms previous findings of NAD processing being functional during early infancy, see [25,29].

- \*\*36. Ravignani A, Sonnweber R-S, Stobbe N, Fitch WT: **Action at a distance: dependency sensitivity in a New World primate.** *Biol Lett* 2013, **9**:20130852–20130852.

Squirrel monkeys (*saimiri sciureus*) went through a habituation-discrimination procedure and showed the ability to recognize whether strings of sine wave tones corresponded to an ABnA rule. A and B consisted of a category of high vs. low tones. Monkeys discriminated the items even when the number of B elements during discrimination exceeded the number of B elements during habituation and also, when the items of the A and B categories were swapped. The study provides a starting point for the investigation of NADs in the auditory domain in primates and should be complemented by studies using different types of stimuli and primate species with different phylogenetic relations to humans and different species-specific communication systems.

37. Newport EL, Hauser MD, Spaepen G, Aslin RN: **Learning at a distance: II. Statistical learning of non-adjacent dependencies in a non-human primate.** *Cogn Psychol* 2004, **49**:85–117.
38. Wilson B, Smith K, Petkov CI: **Mixed-complexity artificial grammar learning in humans and macaque monkeys: evaluating learning strategies.** *Eur J Neurosci* 2015, **41**:568–578.

39. Inoue S, Matsuzawa T: **Acquisition and memory of sequence order in young and adult chimpanzees (Pan troglodytes)**. *Anim Cogn* 2009, **12**:59–69.

\*\*40. Sonnweber R, Ravnani A, Fitch WT: **Non-adjacent visual dependency learning in chimpanzees**. *Anim Cogn* 2015, **18**:733–45.

Chimpanzees were trained to discriminate different types of non-adjacent associations among visual shapes with a 2-alternative-forced-choice discrimination training. The animals seem to be able to learn NADs if they are coded by perceptual similarities and possibly also if they consist of arbitrary relations. Yet, this study does not address whether specific dependencies between arbitrary NADs were learned rather than rules relating visual shape categories to certain sequential positions. The study is the first to successfully test the processing of NADs in humans closest phylogenetic relatives.

41. Endress AD, Carden S, Versace E, Hauser MD: **The apes' edge: positional learning in chimpanzees and humans**. *Anim Cogn* 2010, **13**:483–495.

\*42. Milne AE, Petkov CI, Wilson B: **Auditory and visual sequence learning in humans and monkeys using an artificial grammar learning paradigm**. *Neuroscience* 2017.

Humans and macaque monkeys were tested on sequences of shapes or computer generated sounds from the same finite-state artificial grammar. Both species showed a behavioral sensitivity to adjacent but not non-adjacent dependencies between elements. In both humans and macaques performance was correlated with transitional probabilities between elements. Within each species response patterns were highly similar across modalities, suggesting adjacent sequence processing in different modalities is supported by similar mechanisms that may have been evolutionarily conserved.

43. Ghirlanda S: **Can squirrel monkeys learn an ABnA grammar? A re-evaluation of Ravnani et al. (2013)**. *PeerJ* 2017, **5**:e3806.

44. Newport EL, Aslin RN: **Learning at a distance I. Statistical learning of non-adjacent dependencies**. *Cogn Psychol* 2004, **48**:127–162.

\*\*45. Milne AE, Mueller JL, Männel C, Attaheri A, Friederici AD, Petkov CI: **Evolutionary origins of non-adjacent sequence processing in primate brain potentials**. *Sci Rep* 2016, **6**:36259.

Macaques were exposed to NADs based on spoken syllables in an experiment closely modelled after a study that had been conducted with humans [26]. The primates showed similar electrophysiological brain responses as human infants when presented pitch and NAD deviants during an oddball



paradigm. This study showed, for the first time, a potentially neurophysiological precursor of syntax-related non-adjacent sequencing mechanisms in a close phylogenetic relative.

46. Ueno A, Hirata S, Fuwa K, Sugama K, Kusunoki K, Matsuda G, Fukushima H, Hiraki K, Tomonaga M, Hasegawa T: **Auditory ERPs to Stimulus Deviance in an Awake Chimpanzee (*Pan troglodytes*): Towards Hominid Cognitive Neurosciences.** *PLoS One* 2008, **3**:e1442.
47. Friederici AD: **Evolution of the neural language network.** *Psychon Bull Rev* 2017, **24**:41–47.
48. Lenneberg E: *Biological foundations of language*. New York, Wiley; 1967.
- \*49. Attaheri A, Kikuchi Y, Milne AE, Wilson B, Alter K, Petkov CI: **EEG potentials associated with artificial grammar learning in the primate brain.** *Brain Lang* 2015, **148**:74–80.

Surface-level EEG was recorded as macaques passively listened to sequences of nonsense words that were ordered according to a finite-state artificial grammar. In response to violations of the artificial grammar, the animals showed evidence of an early MMN and P200, as well as a later frontal P500. The study replicated electrophysiological responses that have been observed in humans in response to artificial grammar violations, suggesting that aspects of adjacent sequence processing are likely to have been evolutionarily conserved.

50. Kikuchi Y, Attaheri A, Wilson B, Rhone AE, Nourski K V., Gander PE, Kovach CK, Kawasaki H, Griffiths TD, Howard MA, et al.: **Sequence learning modulates neural responses and oscillatory coupling in human and monkey auditory cortex.** *PLOS Biol* 2017, **15**:e2000219.

### Figure Captions

**Figure 1: Ontogeny: NAD learning in a natural, non-native grammar.** The figure summarizes the experiments reported in [25,26]. A) Stimulus presentation scheme in the learning-test paradigm. Learning phases with correct exemplars (containing NADs) and test phases with both correct and incorrect (containing NAD violations) exemplars are presented in an interleaved procedure. B) Examples of correct and incorrect Italian sentences used in the learning-test paradigm. C) ERP results: 4-month-old infants show a positivity in response to incorrect exemplar sentences compared to correct sentences in the testing phases during passive listening. D) Adults show no ERP differences in response to incorrect and correct sentences in the testing phases during passive listening. E) Adults show a positivity and a subsequent negativity in response to incorrect compared to correct sentences in the testing phases during active listening. F) The arrow symbolizes that the development of cognitive control functions, driven by prefrontal cortex maturation, may underlie the developmental changes of learning-related ERP patterns.

**Figure 2: Ontogeny and phylogeny: NAD learning in an AXB grammar.** The figure summarizes some of the findings reported in [29,45]. A) Stimulus presentation scheme in the oddball paradigm. Frequently presented correct exemplars of the AXB rule (standard stimuli) are interspersed with infrequently presented NAD violations (rule deviants with A elements being followed by the non-predicted B) and pitch violations (pitch deviants with correct NADs, but increased pitch of the B element). B) Macaques show an MMN followed by a later positivity for the pitch violation already during early testing sessions, while the same pattern emerges for the NAD violations during later sessions. C) Three-month-old human infants display an MMN for the pitch and NAD violations. D) Human adults show an MMN for the pitch violations, but no response to the NAD violations. E) The arrow depicts a graded difference across phylogenetic distance and more qualitative difference across ontogenetic distance.

**Conflict of interest statement**

The authors are not subject to any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work.

**Acknowledgements**

This work was supported by the German Research Foundation (project MU 3112/3-1, JLM and project FR 519/20-1, CM), the Max Planck Society (CM), and and Medical Research Council (MRC, U.K.) PhD Studentship (AM).