

Journal of Comparative Psychology

Artificial Grammar Learning in Tamarins (*Saguinus oedipus*) in Varying Stimulus Contexts

Julie J. Neiworth, Justin M. London, Michael J. Flynn, Deborah D. Rupert, Owen Alldritt, and Caleb Hyde

Online First Publication, March 9, 2017. <http://dx.doi.org/10.1037/com0000066>

CITATION

Neiworth, J. J., London, J. M., Flynn, M. J., Rupert, D. D., Alldritt, O., & Hyde, C. (2017, March 9). Artificial Grammar Learning in Tamarins (*Saguinus oedipus*) in Varying Stimulus Contexts. *Journal of Comparative Psychology*. Advance online publication. <http://dx.doi.org/10.1037/com0000066>

Artificial Grammar Learning in Tamarins (*Saguinus oedipus*) in Varying Stimulus Contexts

Julie J. Neiworth, Justin M. London, Michael J. Flynn, Deborah D. Rupert, Owen Alldritt, and Caleb Hyde
Carleton College

The human ability to detect regularities in sound sequences is a fundamental substrate of our language faculty. However, is this an ability exclusive to human language processing, or have we usurped a more general learning mechanism for this purpose, one shared with other species? The current study is an attempt to replicate and extend Hauser, Weiss, and Marcus's (2002) retracted study (2010) of artificial grammar learning in tamarins to determine if tamarins can detect an underlying grammatical structure in a pattern of sounds. Human language consonant–vowel (CV) combinations from Hauser et al.'s original study, newly created tone sequences, and newly created monkey vocalizations made into sequences were used to familiarize tamarins to an AAB or ABB pattern. Tests of novel sounds in each condition were presented that either were consistent with the familiarized pattern or were different from it. Longer looking times toward the sound source (an audio speaker with a specific location in the auditory field) indicated recognition of novelty. Tamarins looked toward the speaker significantly longer with inconsistent human language CV sequences and with inconsistent tone sequences but not when an inconsistent monkey vocalization was presented. Moreover, tamarins showed differential rates of habituation to the different types of sound patterns, with more robust habituation to CV sequences and tone sequences than to monkey call sequences. The implications of these findings for the generality of learning mechanisms for linguistic and nonlinguistic input across species and the importance of testing across various stimuli are discussed.

Keywords: monkeys, grammatical learning, language, music, statistical learning

When speaking a language, humans use a limited set of consonant–vowel (CV) speech sounds, combining them to create words according to grammatical rules and syntax to convey meaning. Because of the complexity of human language and the necessity in language understanding to generalize beyond the specific discrete speech input, researchers and theorists have claimed that human language must include several innate constraints applied to the speech sounds we encounter (Chomsky, 1965; Pinker, 1984). Whether language comprehension is driven by a rule-based set of

hierarchical syntactic structures or emerges from an analysis of the statistical relationships among language elements is currently under debate (Ding, Melloni, Tian, & Poeppel, 2016). Statistics-based models propose that the probabilistic relationships between adjacent words are sufficient for human language comprehension (Bates & Elman, 1996). Others have argued that statistical learning using transitional probabilities cannot reliably segment words but rather relies on constraints from knowledge of phonological structure (Yang, 2004). In fact, there is physiological evidence that low-frequency neural entrainment in humans occurs corresponding with properties such as rhythms in music or natural speech (Zhou, Melloni, Poeppel, & Ding, 2016), and it is hypothesized that neural synchrony at low frequencies with stimuli allows for selective attention to the temporal structure of speech and language (Ding et al., 2016; Ghitza, 2016). One means to examine the prominence of statistics-based models to decipher structures in complex stimuli such as language and music is to test their use by human babies and by other species of animals. The detection of statistical regularities in acoustic input might have evolved in nonhuman animals as well as human animals to recognize patterns of stimuli in their world, and if so, they should be similarly utilized across species and with both communicative and noncommunicative stimuli.

Most nonhuman animal communication comprises vocalizations that convey messages critical to survival, including alarm calls, food-related calls, and calls to signal territorial defense. In many species, there seems to be innate production of these calls, although some training might occur from a tutor at critical developmental periods (ten Cate & Okanoya, 2012). Many nonhuman species' communications show some structure or pattern that has

Julie J. Neiworth, Department of Psychology, Carleton College; Justin M. London, Departments of Music and Cognitive Studies, Carleton College; Michael J. Flynn, Linguistics Program, Carleton College; Deborah D. Rupert, Owen Alldritt, and Caleb Hyde, Department of Psychology, Carleton College.

Deborah D. Rupert is now at the Medical Scientist Training Program, Stony Brook School of Medicine. Owen Alldritt is now at the Department of Philosophy, Emory University.

This research was funded in part by a National Institutes of Health AREA grant (R15HD072571-01) to the Julie J. Neiworth.

The authors thank Sarah Meerts and Julia Strand for editing advice on the manuscript. The authors also recognize the following undergraduate researchers who helped with this project: Guldamlal Kalendar, Catherine Townes, Rae Wood, Deborah Tan, Jordan Palacios, Sofia Levinson, and Evia Zack, Carleton College.

Correspondence concerning this article should be addressed to Julie J. Neiworth, Department of Psychology, Carleton College, One North College Street, Northfield, MN 55057. E-mail: jneiwort@carleton.edu

been characterized as phonological syntax (Marler, 2000) in the sense that there are smaller units of sound used to construct element sequences that have particular statistical regularities in terms of interitem adjacencies. Production of calls by animals is typically limited to unique units (often identified with a sonogram) used to build a unique finite structure with a predictable probabilistic pattern. But does animal communication production aptly describe the means by which animals recognize calls or other incoming stimuli? Studying the ability of animals to perceive acoustic structures of a certain level of complexity is important because it can provide insight into the precursors of language and how linguistic detection has evolved and emerged.

One fundamental generalizable mechanism in language learning is statistical learning, or the detection in a stream of sounds of the high probability events of sounds that typically make up units and words. For example, Saffran, Aslin, and Newport (1996) exposed 8-month old infants to three-syllable CV combinations (e.g., *bidaku*) in a continuous 2-min stream followed by test streams that presented three-syllable words either in a familiar pattern or in a unfamiliar pattern. Infants showed a differential response to novel unfamiliar patterns, indicating that they detected a difference within the pattern of syllables to create “words” and differences between words to demarcate transitions. Saffran, Johnson, Aslin, and Newport (1999) also used tone sequences with 8-month old infants and adult humans and demonstrated abstract pattern detection in both groups that was comparable to that found in the linguistic streams. Similar discriminative behaviors across the domains of language and tones imply a more universal skill for segmentation detected naturally by humans. But can other animals segment statistical patterns of sounds in this way? Hauser, Newport, and Aslin (2001) habituated cotton-top tamarins to speech streams and found that they looked longer toward speakers when the familiar speech patterns changed. Toro and Trobalon (2005) presented CV combinations in a continuous stream to rats that were trained to respond for rewards during the speech stream. When the stream was altered to reflect a different statistical pattern, the rats changed their response pattern, indicating that they detected a change in the pattern of words to nonwords. In sum, the same statistical learning mechanism operates on both linguistic and nonlinguistic stimuli in humans and across various mammals, including rats, primates, and humans. One limitation in humans is created when the statistical pattern is embedded within several patterns, as was done with two different dialects of language streams for 12-month old infants (Gonzales, Gerken, & Gomez, 2015). In that case, infants did not differentiate patterns in the language stream well if it was completely intermixed with a different language stream. So some accumulation of the regularity of the pattern is needed in statistical learning in infants.

A more sophisticated cognitive ability to discern language patterns is to learn that words have to be put in a particular order, and that order can be inferred from examples and then applied such that novel words can be substituted in place of the learned set. Gómez and Gerken (2000) exposed infants to a finite set of “word” strings that convey particular regularities in the construction of “sentences” and thus had a grammar and later tested the infants with completely novel “words” placed in either the habituated grammar or in an ungrammatical form that violated the prior rules. Infants looked longer toward speakers playing ungrammatical sequences, and they generalized their grammar learning to novel

words placed in the proper order. Saffran and colleagues (2008) presented a finite small set of words with a predictive grammatical structure to tamarins, and they reacted with surprise when there were deviations from the predicted structure, but only if the set of stimuli and positioning of stimuli were fixed and no more than one particular item could occupy a position in the grammatical structure. Pigeons were trained to discriminate between several grammatical patterns defined by particular strings of colored letters presented visually all together (Herbranson & Shimp, 2008). They were then tested with novel strings and could detect some differences in grammatical structure, although a detailed analysis of their discrimination revealed that they based their discrimination on subgroups of letter strings common in both the trained set and test sets. Abe and Watanabe (2011) used song elements parsed from Bengalese finch songs and constructed predictive grammatical structures to test whether the Bengalese finches could detect familiar grammatical structures. Bengalese finches changed their own calling pattern when novel strings were presented that deviated from the predictive grammar used in the habituation phase. Upon further analysis, the finches seemed to attend to repeated elements to form discriminations. Tamarins, rats, pigeons, and finches can discriminate in a limited way grammatical sequences, but their discrimination is typically based on noticing repeated units within the structure or remembering single items in particular positions or based on a unique beginning or ending to the structure. Their responding reveals a lack of flexible or more abstract understanding of grammar structure.

A detection ability more complex than judging changes in a continuous auditory stream (statistical learning) but not elevated to the level of recognizing that words require a particular order is the ability to perceive abstract patterns (i.e., ABB) that, once learned, are impervious to switches in the elements used. This ability, which applies statistical learning to extract abstract rules about patterns, is tested within an artificial grammar learning (AGL) paradigm. In these studies, human subjects are first exposed to particular CV combinations that are structured in a particular order. Later, subjects are exposed to novel examples of CV combinations that either match the prior familiar algorithm or pattern or deviate from the prior pattern. Marcus, Vijayan, Bandi Rao, and Vishton (1999) constructed three-syllable CV combinations that either matched an ABA pattern (i.e., *ga-ti-ga*, *li-na-li*, *ta-la-ta*) or an ABB pattern (i.e., *ga-ti-ti*, *li-na-na*, *ta-la-la*) and habituated 7-month-old infants to them. When novel three-syllable patterns were presented, infants looked longer at the source speaker when the pattern deviated from the familiar one (from ABA to ABB, or vice versa) than when they heard novel strings that matched the familiar pattern (ABA followed by novel ABA). Of course, these patterns allowed infants to detect the difference based on a repeated syllable (as in ABB) as opposed to a nonrepeating pattern (ABA), which can be construed as a simpler detection task relying upon duplication and not evidence of rule-based patterns. Marcus et al. (1999) controlled for this by using two repeating patterns, AAB or ABB, in a second test with 7 month olds who subsequently looked longer when a different novel repeating pattern was presented (i.e., a switch from an AAB pattern to an ABB pattern). Hauser, Weiss, and Marcus (2002) used the AGL paradigm and habituated tamarins to either AAB or ABB patterns of CV combinations, and they reported that the tamarins looked longer when they changed the pattern (i.e., from AAB to ABB). Their study

suggested that tamarins have an ability to derive abstract rules; however, it was retracted when it was found that “the data do not support the reported findings” (Hauser, Weiss, & Marcus, 2010). In another experiment in which the patterns ABB or AAB with three-syllable word construction were used, rats did not extract the different rules from the patterns because their response patterns were similar with changes between AAB and ABB (Toro & Trobalon, 2005). Corballis (2009) found that rats discriminated three-item patterns based on the repetition of units within the pattern or by learning the first or last element of the three-item set. In a similar study, Murphy, Mondragon, and Murphy (2008) trained rats in a classical conditioning paradigm to anticipate food following a pattern of light and darkness that was of a structure *XYX*, *XXY*, or *YXX*. In a test, deviations from the original pattern induced rats to suppress anticipatory responses. Although this and other studies similar to it show that rats can be sensitive to regular patterns paired with outcomes, they could be noting very simple elements of duplication or nonduplication (i.e., two flashes of light in a row, or an alternating pattern), which does not indicate that they acquired an abstract rule of the pattern. A transfer test using auditory stimuli was also conducted, but it suffered the confound that repeated presentations and not an overall pattern could have been the cue for the rats. In another study using ABA or BAB patterns constructed of song elements (van Heijningen, Chen, van Laatum, van der Hulst, & ten Cate, 2013), finches could learn to differentiate the patterns in a go/no-go response paradigm, but only one of the eight finches studied could generalize the discrimination to novel sets in which one element was replaced by a novel element. However, in a more recent study by Spierings and ten Cate (2016), budgerigars (a parrot species) and zebra finches were trained to discriminate *XYX* or *XXY* patterns in a go/no-go paradigm, and budgerigars exhibited transfer of learning to some novel sets, even with unfamiliar items, although responding was quite low in the novel conditions. Finches did not demonstrate this flexibility of applying a learned pattern to novel sets. Thus, using the AGL paradigm, only humans (infants and adults) and possibly budgerigars, with the one recent study, demonstrated discrimination between different abstract patterns that supersede changes in the actual units on which the patterns are built. Tamarins were reported to demonstrate this as well, but the data from Hauser et al. (2002) were later reported to not support the published findings (Hauser et al., 2010). The other tests of nonhuman animals demonstrate that the memorization of fixed pairs or the use of repetition to form a discrimination and not the detection of a pattern more abstractly applied.

Another critical difference in the mechanisms used by humans to acquire language might be constraints on their use of particular types of input. Saffran (2002) found that adult human learners and children can detect statistical properties of linguistic input such that some word categories such as *the* or *a* can be used to predict phrase units and thus detect typical dependencies in language. Saffran found that, in humans, the detection of phrase structure by the use of cues can be demonstrated in artificial language learning and occurs with nonlinguistic sounds (i.e., computer alert noises, drums, and bells) as well. In contrast, Saffran discovered that subjects could not use phrase dependencies when visual stimuli were presented unless they were presented simultaneously rather than sequentially, the means by which all of the auditory stimuli were presented. Marcus, Fernandes, and Johnson (2007) found an

inability by 7.5-month-old infants to generalize an habituated abstract pattern of tones or animal call sequences to novel examples consistent with that pattern, but the infants exposed to sung syllables in ABB sequences could generalize the abstract pattern to consistent novel sequences of human-sung syllables. In fact, if the infants were habituated to sequences of spoken syllables, then they could then generalize the pattern they learned to sequences of tones and animal sounds consistent with the CV pattern, but only when they had been pretrained to habituate to the ABA or ABB patterns in human spoken language. Likewise, Ferguson and Lew-Williams (2016) found that 7-month-old infants could generalize their habituation of ABB sequences using sine-wave tone sequences, and looked longer when the sequence was violated (ABA), but only if they had been preexposed to videos in which people were conversing and one of the two people had the same sine-wave tones dubbed over her speech such that the tones appeared to have communicative meaning. Starting only with tones in their pure form, infants did not extract a pattern rule that could be applied to language-like stimuli. This specificity of input to generate abstract rule regularities has also shown failure by 7.5 month olds to abstract rules from sign-language-like gestures (Rabagliati, Senghas, Johnson, & Marcus, 2012). In sum, human and nonhuman animals can detect statistical patterns in auditory sounds, including human CV sequences. Even as infants, humans seem unique in tolerating novel words substituted in the set, whereas other tested animals, including rats, pigeons, finches of various strains, and tamarins, tend to memorize subgroups of patterns or individual units to discriminate grammatical structures in a limited way. There seems to be a limitation on human infants and adults in applying learned algebraic rules such as ABB or AAB to novel patterns in that generalization of consistent patterns does not occur with visual sequential input, with animal sounds, and with tones. The exception seems to be if the test occurs after a training block with human CV sequences, either spoken or sung, and then generalization of the pattern can occur to the other acoustic stimuli.

Tamarins may demonstrate more sophisticated use of general learning mechanisms to acoustic input, but proper tests have not been conducted or were conducted but then were retracted. As a monkey species with great evolutionary distance from humans, it is important to determine the kinds of detection and generalization around statistical learning and AGL that occur for them, and whether they also show constraints on applying rules they have learned to particular acoustic input. Tamarins demonstrated a similar ability to humans and other animals to detect pattern changes in previous studies, but it is not clear whether they can apply an abstract pattern conveyed in human CV sequences to novel consistent patterns. Moreover, AGL has not been studied using different sound types in tamarins in past studies. A test using the AGL paradigm and using human vocal sounds, tone sequences, and monkey calls in the current study will help to determine whether tamarins can detect abstract patterns generally and can do so independent of sound type.

The current study attempts to replicate the Hauser et al. (2002) study of AGL in tamarins. It is not a strict replication, although the original CV combinations were obtained from Hauser for use in the present study. Habituation criterion was defined in Hauser’s study as a measured decrease in looks in several consecutive behaviors emitted by individual monkeys toward a sound source.

This definition operationalized habituation in terms of a fleeting change in behavior that was judged while it happened by researchers conducting the study. Instead, in this study, a total of 40 exposures of each sound type was presented across 5 sessions/days. This exposure rate exceeds that used to study infants in the original studies by Marcus et al. (1999) and matches the maximum needed to generate habituated responses by tamarins in the former Hauser study. Post hoc analyses of looking per trial were conducted under conditions that made coders blind to the sound type and the progression. The present study uses the human vocal CV combinations in one condition, but it also uses tones to construct similar AAB or ABB patterns in another condition, and in a third condition it uses monkey vocalizations parsed into small units and presented as AAB or ABB patterns. These sound types have been used to investigate the universality of applying statistical models by human infants. Novelty was defined in the tests by introducing novel CVs, tones, and calls and by presenting them in consistent (same) or inconsistent (different) patterns. In the case of tone sequences, both test stimuli also presented a different contour (an ascending pattern) than the habituated pattern (a descending one) so that simple transposition could not account for detection of the consistent pattern in the test; rather, a recognition of ABB or AAB was necessary.

Method

Subjects

The subjects were 16 adult-aged cotton-top tamarins, (*Saguinus oedipus*) housed in pairs or triplet groups in three different monkey colony rooms in the animal facility at Carleton College. There were 10 females and 6 males in the original CV condition, but in one auditory condition (CV combinations) a female subject was not visible in the video recording through both test trials; therefore, only 15 subjects' data were analyzed. Three females died over the course of the study, leaving data from 13 monkeys in the tones condition and the monkey call condition. The ratio of females to males in the final two conditions was 7:6. None of the subjects had participated in an auditory discrimination experiment before this study, although they had heard humans talking and monkey calls of relatives and unfamiliar tamarins in the colony during their lifetimes. The age range of the monkeys was 5 (Egret) to 19 (Quince) years, with adult onset occurring around 21 months. All subjects had been adults for at least 3 years before the study began; therefore, the age range only captures a difference in adult ages, not a developmental difference.

Because the tamarins were tested in pairs or triplets in their home cages, it was important to determine whether tamarins' looks in the various tests could be independently treated. Correlation between looking times between pairs or among triplets in testing revealed a strong and positive relationship among cagemates in CV sequence testing, ($r = .856, t(12) = 5.24, p < .01$), somewhat weaker but still positive correlation among cagemates in tones testing ($r = .54, t(12) = 2.02, p = .07$), and relatively uncorrelated reactions during monkey calls testing ($r = .12, t(12) = 0.37, p = .71$). Because of the heavy influence among cagemates in looking times, an average score for each pair or triplet was used for look times in all analyses. Thus, the total n in each condition was reduced from 13–15 individual monkeys, depending upon condi-

tion, to 8, 6, and 7 caged groups in the CV test, tones test, and monkey calls test, respectively.

The initial group, described earlier, was all exposed to the same order of auditory conditions (CV sequences, tones, monkey calls, respectively). A second group of five monkeys was recently acquired and used to run the study in a different order (tones, monkey calls, CV sequences, respectively). There were two females (Haagen Daz, Wisteria) and three males (Vireo, Echinecia, Yogi) with ages ranging from 12 to 17 years. Their looking times were collected and averaged per cage, which yielded three caged groups (Yogi and Haagen Daz were singly housed).

The monkeys were monkey-family reared in laboratory settings and had been socially housed in pairs in seven different $0.85 \times 1.50 \times 2.30$ m cages, with the cages visually separated by opaque sheets. The subjects were on a 12-h light–dark cycle and had free access to water. All animals were maintained on a complete diet consisting of a yogurt and applesauce breakfast; a lunch of Zupreem Marmoset chow, fruits, and vegetables; and a protein snack (e.g., eggs, hamburger, mealworms) each day. The protocol and care of the monkeys were approved by the Institutional Animal Care and Use Committee, the monkeys were inspected regularly by the U.S. Department of Agriculture, and their care consistently met animal welfare assurances.

Materials

There were three conditions involving sound patterns: CV sequences, tones sequences, and monkey call sequences. Each sequence consisted of sound units that were either CV combinations, single tones, or a monkey call parsed into a single unit. The characteristics of one type of sequence and of the units that comprise the sequences are given in Figure 1.

All CV combinations were obtained from Hauser and were the female “Judy” voice used in the original study (Hauser et al., 2002). Each CV unit was a CV combination of 0.5- to 0.7-sec duration. The three-unit CV sequences were built from these CV combinations using Audacity 2.0.4 at Carleton College. CV sequences consisted of three elements in either an AAB or ABB pattern with a 1-sec inter-onset interval. There was a delay of 0.3–0.5 sec between CV units spoken, depending upon the natural decay of the CV spoken. There was a 2.5-sec interval between successive three-CV sequences (see Figure 1). The consonants and vowels used to build the CV combinations in the familiarized set included plosive consonants (p, b, d, k, t) and nasal consonants (m, n), together with vowels that were either close front (i) or close back (u). Thus, for example, the three-CV combinations that made up the AAB patterns included “didibu,” “pupuki,” “mumuni,” “pipigu,” “bibidu,” “nunumi,” “duduki,” “pipitu,” “titiku,” and “kukudi.” The ABB sequences used the same consonants and vowels except in the ABB pattern (e.g., “dibubu” for the first sequence). The Test CV combinations were constructed with novel glide consonants (w, r, j) or liquid consonants (l) and a central open vowel (a) or a closed front vowel (i). The Test CV sequences were “lalari” for the AAB sequence and “wajiji” for the ABB sequence. Each three-CV sequence was played 4 times to form a trial. Within a trial, each three-CV sequence was started on a 5.0-sec mark; therefore, the first sequence was started at 0 sec, the second sequence at 5.0 sec, the third sequence at 10.0 sec, and the fourth at 15 sec. There was a delay of approximately 2.5 sec placed

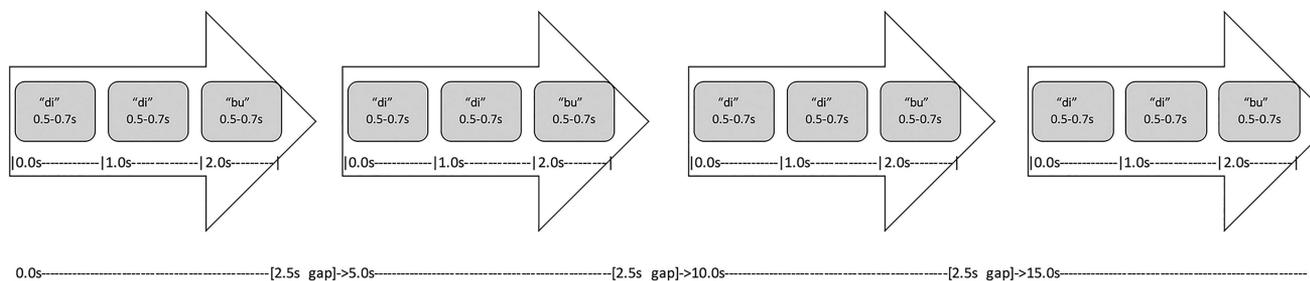


Figure 1. Graphic representation of an AAB sequence, presented 4 times, in a single trial. Sounds and sound length are presented in arrows. Times directly below graphics show the start time of each element and the gap between three-item sequences. Times shown at bottom indicate the beginning of each three-item sequence.

to separate three-CV sequences. Each CV trial lasted a total of 17–20 sec in duration.

For the tone sequences, tones were selected from frequencies from C5 to C7 (523.251–2,093 Hz) to match the frequency range of monkey calls of this species, which is rather high. Each note was generated using the “piano” instrument in Garage Band software (6.0, Apple, Inc. Cupertino, CA) and then edited in Audacity. All other characteristics matched the CV sequences, including the duration of tones, the gaps between tones, the three-unit sequence, the gap between sequences, and the presentation of four sequences to define a trial. Tone durations were between 0.5 and 0.7 sec, which matched the duration of the spoken CV combinations from Hauser et al. (2002). The XXY¹ patterns used in training were CCE, AAD, EEB, FFA, DDG, AAE, CCF, AAC, GGB, and EEB—thus two tones and then a third tone at least 3 and no more than 10 semitones higher. The XYY patterns used the same note pairs as in the XXY pattern; thus, for example, the first sequence was CEE. All familiarized sequences involved an ascending melodic interval. The test sequences used a descending interval; thus, the syntactic generalization to be learned was the pattern (XXY vs. XYY) and is independent of melodic direction. The same test sequences were used for both groups of habituated monkeys. They were GDD (XYY pattern) and FFC (XXY pattern), with both sets of tones in octaves not used for those particular tones in the habituated set and both descending in pattern while all habituated patterns were rising or ascending.

The monkey call sequences were constructed from tamarins’ calls recorded in 10- to 15-min sessions by two undergraduate researchers during feeding times and during a time of stress, when a confederate student donned a full face mask and marched around the colony rooms to solicit barks and warning calls. The streaming calls were then downloaded to Audacity and edited into 0.5- to 0.7-sec monkey call units. Attempts were made to parse at natural communicative boundaries based on readily identifiable calls, barks, chirps, and screeches used to indicate search for other monkeys, warning calls, and food-related calls. The monkey calls used included four different “long” calls that consisted of ascending continuous vocal sound, four different monkey barks that were vocal calls of shorter bursts, two screeches evoked before running, and several different chirping calls solicited by different food items. These were coded and mixed such that AAB exposure included, for example, Bark1-Bark1-Chirp1 or Bark2-Bark2-Screech1. ABB sequences used the same call units matching the appropriate sequencing; for example, the first sequence was

Bark1-Chirp1-Chirp1. The test call sequences were composed of a High-Low-High (HLH) call, a double chirp (DC1), a long call (LC1), and a double-bark (DB1) which had not been used during habituation. They were HLH-DB1-DB1 for the ABB pattern and LC1-LC1-DC1 for the AAB pattern. The structure of the units, sequences, and trials matched the structure used for CV sequences and note sequences (see Figure 1).

Half of the subjects heard an AAB sequence, played 4 times per trial, and half of the subjects heard an ABB sequence, played 4 times per trial. All sequences were presented using Audacity through an Apple MacBook Pro with external speakers.

Procedure

This procedure is described in Neiwirth (2013) for an analysis of food-eating behavior during the trials. Subjects were exposed either to the AAB pattern or the ABB pattern during familiarization. Whether they actually habituated to AAB or to ABB varied across the three auditory conditions (CVs, tones, or monkey calls). Familiarization consisted of five sessions occurring across 5 consecutive days. In each session, two trials lasting 17–20 sec each in duration were presented. Each trial consisted of four repetitions of a single auditory sequence matching a particular pattern at the same volume. A total of 40 exposures of the pattern occurred in familiarization across the five sessions, with a total of 10 unique three-unit sequences used per auditory condition.

Before each trial began, undergraduate researchers placed digital cameras on tripods in front of each cage in which a pair or triplet of monkeys was housed. The primary investigator (PI) moved the laptop computer and speakers on a small cart behind the cages to a corner of the room. On her cue, the undergraduate researchers would simultaneously enter the cages and deposit several Frosted Cheerios in each food bowl and exit the cage. They would immediately stand behind their cameras and begin recording. The PI would then play the first trial, which consisted of four presentations of the same ABB or AAB sequence, depending upon the assignment for exposure for the monkeys in the room. Each trial lasted approximately 17–20 sec. The camera operators would vary the angle of focus of the camera to capture the faces and heads of the tamarins while the sounds were being played. Because

¹ A switch was made to use XXY and XYY rather than AAB and ABB in the description of tone sequences to avoid confusion between sequence structure and the musical pitches used in their construction.

the tamarins often positioned themselves in front of the bowl with cheerios, it was easy to locate both tamarins in the cage and record their head and body orientation toward the speakers throughout the auditory trial. The researchers would then repeat this process for a second trial to complete each session.

Each pair of subjects was exposed to the training pattern (either ABB or AAB) for five consecutive sessions, with two trials presented per session. In the fifth session, two familiarization trials were presented, followed by two test trials, each composed of four exposures of a novel sequence. One test trial labeled Novel Different presented novel sound units within the same sound type (CVs, tones, or monkey calls), but the pattern was different (i.e., ABB if the subjects had habituated to AAB). One trial called Novel Same presented novel sounds from the same sound type and in the same pattern that was familiarized (AAB if the subjects had been exposed to AAB). It is important to note that each condition only used two test units (e.g., for the CV test, one was “wajiji” repeated four times and one was “lalari” repeated four times), but the design forced a different interpretation on these two stimuli by the subjects if they were applying an abstract pattern rule to the new stimuli. Those exposed to ABB sequences should regard “lalari” as both new and a different pattern whereas those exposed to AAB sequences should regard “lalari” as new but having the same pattern structure. Thus, for example, the subjects’ look rates should differ to “lalari” depending upon whether it constituted a novel pattern to them or not. The same was true for the two novel stimuli in testing in the tone sequence condition and the monkey call condition. Whether the Novel Same sounds were presented first or presented second in testing were also counterbalanced across the three auditory conditions for each pair or triplet of tamarins. The same data, look rates as indicated from video of body and head posture toward the speakers, were recorded in the test trials as well as the habituation trials.

Video data were downloaded in iMovie HD version 6 (Apple, Inc, Cupertino, CA), which allows for manually scanning clips at 30 frames/sec with no auditory sound on. The researchers’ coding could not differentiate which condition was being coded in iMovie because the auditory portion was muted (CVs, tones, or monkey call). Coders were aware or could see on the video where the speakers were placed in the room; thus, the angle of viewing for a tamarin to look at the audio speaker was clear. Trial order was preserved in downloaded digital videos; thus, the two test trials were the last trials on each digital video, but without audio, the coders could not determine which of the two test trials was Novel Same or Novel Different. Thus, the researchers who were coding could not tell which condition they were coding nor which test trial they were coding (new pattern or habituated pattern) to prevent experimenter bias. One primary researcher coded the digital videos for the monkeys in one condition, with a second researcher coding all of two monkeys’ look rates as well so that interrater reliability could be checked. Look rates were entered in an analysis as milliseconds using a formula that converted 30 frames/sec into millisecond timing, $ms = \text{frames of looking toward speaker} * 1,000/30$. Across the three conditions, interrater reliability was determined using Cronbach’s α for two different raters per monkey per trial using millisecond look rates, and the analysis reflected highly correlated coding that varied from $+0.72$ for monkey calls to $+0.89$ for look rates with tone sequences to $+0.98$ for look rates with CV sequences. All researchers’ coding videos were required

to track a single monkey frame by frame through all 10 familiarization trials as well as the 2 test trials and to use head and body orientation as an index of looking. Tamarins have flat faces and very few degrees of pupil movement possible; thus, they turn both head and body toward an item to which they are attending.

One smaller group of monkeys ($n = 5$) was exposed to the same procedure and data collection as described here but with the auditory conditions presented in a different order (tones, monkey calls, CV sequences). The first large group was presented with CV sequences first because that was the test of interest to confirm or disconfirm Hauser et al.’s experiment. Any residual effects from prior exposures should more quickly habituate tamarins to other sound types and patterns, and this clearly did not happen in the original group because the tamarins showed greater initial arousal to second and third auditory sets (tones and monkey calls, see *Results*) and even a lack of habituation to the last set. Fatigue or practice would generate more habituation at a quicker pace as tamarins were exposed to more trials of patterns. To control for this, a small group of monkeys without any experience with auditory stimuli were acquired and used to test a different order of conditions. Their analyses are presented separately at the end of the *Results* section.

Results

An examination of looking times across the 40 exposures to each stimulus type shows whether habituation by the pairs/triplets of tamarins occurred. Figure 2 displays averages of the tamarins’ look rates in blocks of two trials (which constituted single sessions) during the familiarization phase. The best-fitting line for look rates over the 10 trials (40 exposures) to the CV sequences had a negative slope, $y = -753.05x + 5,708.3$, $R^2 = .75$. A paired-sample t test comparing look rates toward the speakers on the first trial of habituation to CVs ($M = 3,052$ msec) to the last

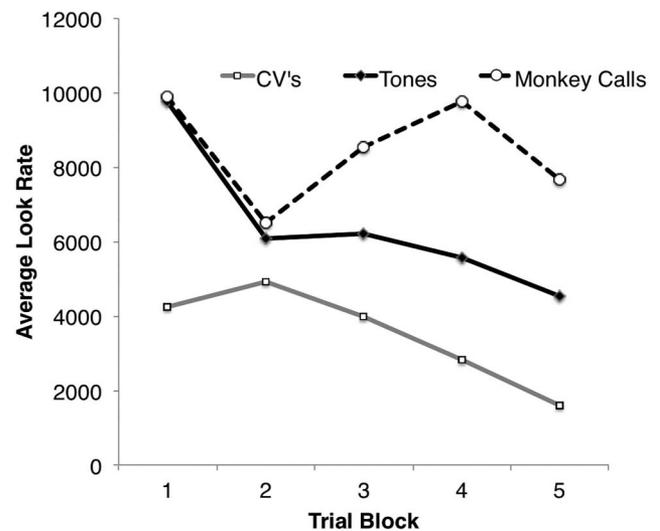


Figure 2. Averages across all of the tamarins’ look rates in blocks of two trials (which constituted single sessions or days) during the familiarization phase, with CV sequences shown first, tone sequences second, and monkey call sequences third. CV = consonant–vowel.

trial ($M = 632$ msec) decreased significantly, $t(7) = 2.37, p = .05$. Moreover, the looking time was 21% of the original average looking time. Often developmental studies with infants use a cutoff of either 75% of the original or 50% of the original as their definition of habituation. By negative slope, significantly lower looking time, and a low percentage of the original looking times, tamarins were habituated to the CV sequences.

There was also habituation to the tones, as demonstrated by a significantly lower look rate to the last trial of tone sequences ($M = 3,499$ msec) as compared with look rates on the first trial ($M = 10,338$ msec; $t(5) = 4.17, p < .01$). The best-fitting line to looking times across the tone sequence trials had a negative slope, $y = -1,074.7x + 9753.9, R^2 = .73$, and the final looking time was 34% of the original duration of looking and lower than criteria used in developmental literature. In contrast, the best-fitting line for the look rate data to the monkey call sequences per session had a slightly positive slope with a very low amount of variance accounted for, $y = 68.99x + 8,345.4, R^2 = .007$. A paired t test comparing the first trial ($M = 10,114$ msec looking time) to the last trial looking times ($M = 7,093$ msec) to monkey calls was not significant, $t(6) = 1.41, p = .21$, and they remained looking at 70% of their original mean looking time.

Did the tamarins' looking times increase appreciably with novel sequences, especially those that also introduced a different pattern? A mixed-model analysis of variance (ANOVA) was conducted comparing looking times across auditory condition (CV, tones, or monkey calls), trial type (last familiarization trial, test Novel Same pattern, test Novel Different pattern), and the fixed factor of initial pattern type (AAB vs. ABB). A significant effect was found among the three trial types ($F(2, 30) = 10.07, p < .01$, effect size = 0.40^2) and between the three auditory conditions ($F(2, 15) = 8.17, p < .01$, effect size = 0.52). Further testing by paired-sample comparisons of the three auditory conditions indicated that the tamarins looked overall for a significantly shorter period of time toward the sound source when CVs were played ($M = 2,363$ msec) as compared with tones ($M = 4,906$ msec, $p = .04$) and monkey calls ($M = 6,808$ msec, $p < .01$). There was also a significant interaction of Trial Type \times Auditory Condition, $F(4, 30) = 4.83, p = .02$, effect size = 0.39 , suggesting that reactions to the different tests did not occur at similar rates across the three auditory conditions. Looking times between the two initial familiarized patterns (AAB vs. ABB) were not significantly different, $F(1, 15) = 0.003, p = .96$, indicating that one particular pattern was not more attention-getting than the other. There were no significant differences across other interactions. In sum, tamarins looked differently toward the different sound sequences (CVs, tones, and monkey calls) and they habituated to and reacted differently to tests of these sequences, as evidenced by a significant interaction of trial type by auditory condition.

To determine whether the tamarins looked longer at a novel sequence in a different pattern than at a novel sequence presented in the same familiar/exposed pattern, paired-sample t tests for each of the three auditory conditions compared looking times to the last familiarized trial to the looking times to the two test trials (Novel Same pattern vs. Novel Different pattern). For the CV sequences, tamarin pairs and triplets (with one individually housed tamarin coded) looked significantly longer at the novel sequence in a different pattern ($M = 3,674$ msec) than at a novel sequence presented in the same pattern ($M = 1,979$ msec, $t(7) = -3.38, p =$

.01). The tamarin groups did not look significantly longer at the novel sequence in the same pattern compared with their looking times to the last exposure ($M = 632$ msec; Novel Same, $t(7) = -1.94, p = .09$). They did look significantly longer at the novel sequence presented in a different pattern than they looked at the last familiarized trial, $t(7) = -3.66, p < .01$. When novel tone sequences were compared, tamarin pairs (one triplet) looked longer toward novel tone sequences in a different pattern ($M = 7,329$ msec) than they did to novel tone sequences in the same familiar pattern ($M = 3,890$ msec, $t(5) = -2.46, p = .057$), but this was a trend and not significant. Moreover, differences between the looking times from the last familiarized trial ($M = 3,499$ msec) to the novel sequence in a different pattern was also longer and a trend, $t(5) = -2.02, p = .10$, but no difference existed between looking toward the sound source during the last familiarized trial and looking during the novel sequence presented in the same pattern ($p = .64$). For monkey call sequences, there were no differences between looking times to the novel sequence presented in a different pattern ($M = 6,752$ msec) compared with the novel sequence that was presented in the same pattern ($M = 5,987$ msec, $p = .37$). There were no differences between the average looking times to the last familiarized trial ($M = 7,093$ msec) and either of the two novel test patterns ($p = .51$ and $p = .74$ for a comparison to Novel Same pattern and Novel Different pattern, respectively). Figure 3 depicts average looking times by cage/group to the last familiarized trial and to each of the two novel test patterns (novel sequence, same pattern or novel sequence, different pattern) across the three auditory conditions.

The two test stimuli within each auditory condition were also examined to see if one simply generated more looking time than the other, independent of the context in which it was presented (as a novel pattern or as a familiar pattern stimulus) by paired-sample t tests. There were no significant differences in look rates to the test stimuli themselves independent of pattern context. For example, the mean look rate to "wajiji" across all tamarins was 2,534 msec and the mean to "lalari" was 2,444 msec, with no difference between these two CV test sequences independent of context, $t(7) = 0.082, p = .94$. It was only when either of these CV sequences was presented as a Novel Different pattern that a significantly longer look rate emerged.

A second group of tamarins newly acquired in the laboratory ($n = 5$) was exposed to the auditory conditions in a different order (tones, monkey calls, CV sequences). One tamarin died 1 week into the study of a chronic health issue; thus, four tamarins participated, with two in a pair and the other two singly tested. Their results mirror the original group's reactions without any noticed effect produced by the different order. The tone sequences were their first familiarized set, which generated a best-fitting line that was negative, $y = -1,168.3x + 9,000, R^2 = .77$ (see Figure 4). Their second familiarized set was comprised of monkey call sequences, and that generated a slightly negative best-fitting line that showed very little variance accounted for, $y = -513.33x + 10,866, R^2 = .14$. The final familiarized set comprised CV sequences, which generated the shortest looking times but also a negatively sloped best-fitting line, $y = -1,298.7x + 6,718.1, R^2 = .77$.

² Effect size is defined by partial η^2 , generated in SPSS in the ANOVA.

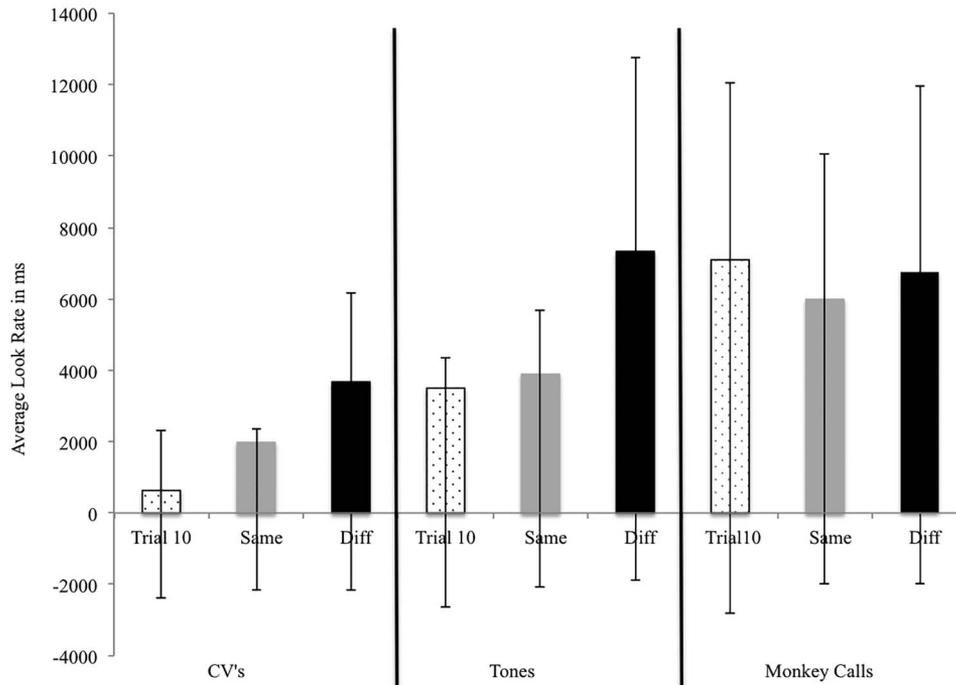


Figure 3. Average looking times and 95% confidence intervals for the last trial of familiarization (Trial 10), the novel sequence in the same pattern, and the novel sequence in the different pattern for each sound type (CVs, tones, and monkey calls). CV = consonant–vowel.

To determine if the second group also reacted more to the novel sequences in different patterns, a mixed-model ANOVA comparing trial types (last familiarized trial, Novel Same, Novel Different) and auditory condition (tones, monkey calls, CVs) produced a significant interaction effect of Trial Type \times Condition, $F(4, 12) = 3.56$, $p = .04$, and a significant main effect of auditory condition ($F(2, 6) = 15.25$, $p < .01$; see Figure 5). Looking times to monkey calls in general ($M = 9,386$ msec) were significantly longer than were looking times during CV sequences ($M = 988.15$ msec, $p < .01$) and to tones ($M = 4,064$ msec, $p = .01$). There was a trend difference ($p = .09$) indicating shorter looking times to CV sequences than to tones, but this difference was not significant in the small control group used. In the CV sequence test, paired-sample t tests revealed significantly longer looking times to novel sequences in a different pattern ($M = 1,922$ msec) than to novel sequences in the same pattern (697 msec, $t(2) = -5.27$, $p = .03$). Moreover, looking times to novel tone sequences presented in the same pattern were also significantly shorter ($M = 1,566$ msec) than were looking times to novel tone sequences presented in a different pattern ($M = 7,077$ msec, $t(2) = -4.49$, $p = .05$). There was not a significant difference between average looking times to novel monkey calls presented in the same pattern ($M = 12,782$ msec) as compared with a different pattern ($M = 8,144$ msec).

Discussion

This study found that tamarins detected a change of pattern in human spoken CV sequences after being familiarized through repeated exposure to a particular pattern, and this was supported by significantly longer looking times by tamarin pairs and triplets

toward novel sequences presented with an unfamiliar pattern as compared with novel sequences presented with the familiar pattern. This occurred independent of the placement of the CV sequence condition within the order of auditory conditions and was not a result of fatigue or any artifact of the study's design. Thus, the findings of the original study by Hauser et al. (2002) specific to human CV sequences were confirmed. However, an exploration of different auditory contexts allowed for a deeper understanding of the tamarins' reactions to patterns. The effect within the tone sequences was similar, but it generated a trend difference in longer looking times toward novel sequences presented in different patterns in the first large group and a significant difference with longer looking times to the novel pattern as compared with the familiarized pattern in the second smaller group. The effect within the tone sequences was not caused by transposition of a pattern or melody to a different key, for the test tone sequences were all constructed with a different contour, an ascending sequence, rather than a descending one and with initial pitches randomized within the C-Major Diatonic Scale. Thus, the significant pitch cue was not based on any memory for absolute pitch but rather a sensitivity to relative pitch specifically, in which the pitch change occurs in the sequence (see Honing, ten Cate, Peretz, & Trehub, 2015). Despite the novelty of the relative pitch change, the monkeys still reacted to novel presentations in the familiarized pattern with less attention. Finally, the tamarins' pattern detection and reaction to new patterns did not surface when monkey calls were used. There were significantly longer looking times generated throughout the familiarization phase to monkey call sequences, although they were presented using the same pattern for each tamarin. There was no

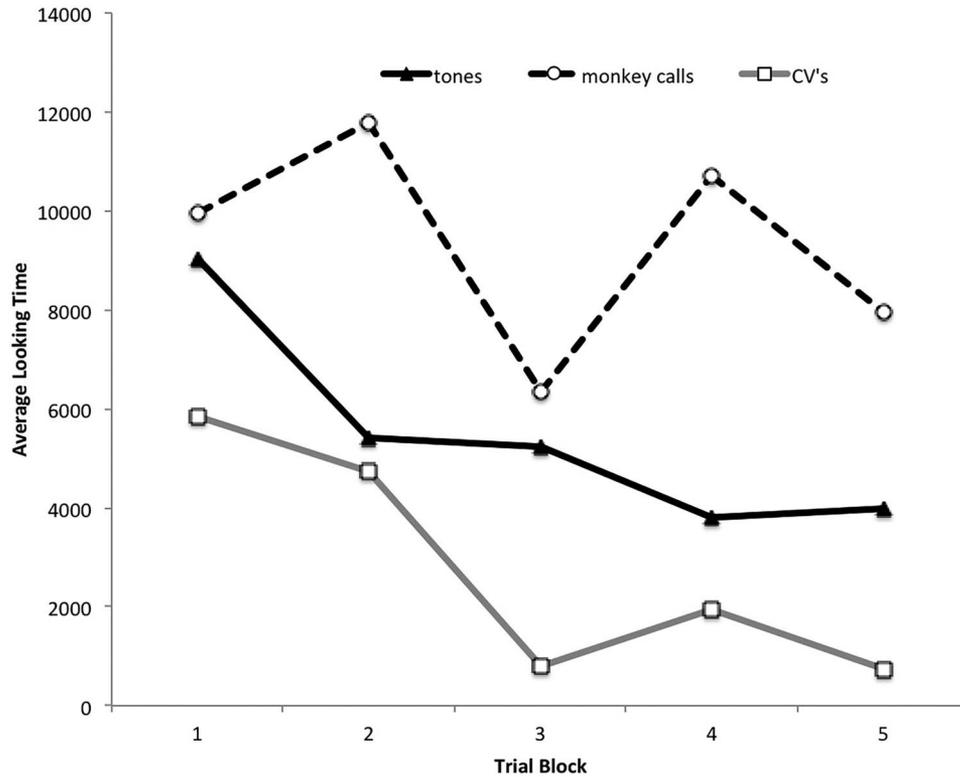


Figure 4. Average looking time in blocks of two trials during familiarization in a different order, with tone sequences first, monkey calls second, and CV sequences last. CV = consonant-vowel.

attentional “savings” toward a novel monkey call sequence presented in the same pattern in the test. Rather, monkeys looked at similarly long rates to novel monkey call sequences whether they were presented in the same or a different pattern in the test. The results as a whole suggest a common mechanism among tamarins for detecting patterns within certain auditory domains, which include human CV sequences and tone sequences. These results are generally similar to the reactions by human infants in past research, although human infants seem biased toward detecting differences in pattern across changes in human vocal sounds as compared with artificial stimuli or animal calls. The results are not due to a discrimination between stimuli based on duplication or alternation within the sets, which can be a discriminative cue for rats and finches in prior studies. In this study, both habituated and novel patterns contain duplication of sounds, and the monkeys still reacted to novel sequences with a different pattern more than to novel sequences in the same pattern if human vocal sounds or tones were used.

It is important to note that detection of the pattern change did not occur to novel patterns of monkey calls. This seems counter-intuitive if a purpose of pattern detection might be to recognize changes in the structure of language or communicative input, for in this case the monkeys did not react more to changes in patterns in their own vocalizations whereas they could detect changes in patterns in other sonic domains (CVs and tones).

The findings suggest that statistically regular patterns were noticed by the monkeys if the monkeys had shown some degree

of habituation to the regular patterns provided in that sound type but not if the patterns perceived during habituation generated high rates of looking time and arousal. It is important to note that the tamarins did not habituate to all three types of sounds. Human CV sequences and tone sequences generate the most robust habituation with significant decreases in looking times across the 40 exposures whereas the monkey calls generate no significant decrease or change in looking times over 40 exposures. For tamarins, noticing the change in pattern depended upon the units of sound becoming less salient by virtue of exposure to the monkeys. Monkey calls constructed to form arbitrary three-unit sequences created strong arousal reactions in the monkeys; in fact, they reacted to their meaning quite often by barking back, running in the cage, or staring at the speakers, depending upon the aberrant combinations. In the face of this kind of high level of unwavering attention during the familiarization phase, the monkeys did not differentiate a change in the overall pattern of presentation using novel monkey call sequences, most likely because they seemed to react to different individual monkey call units presented within the sequences throughout training and testing. In this case, the units and their adjacencies may have defined unexpected combinations such that an overarching pattern did not draw attention. Or it may be that individual units carry more significance and salience in animal communication than does the pattern within which those units are presented. In fact, in the first group with a fixed order of presentation, the monkey call sequences were

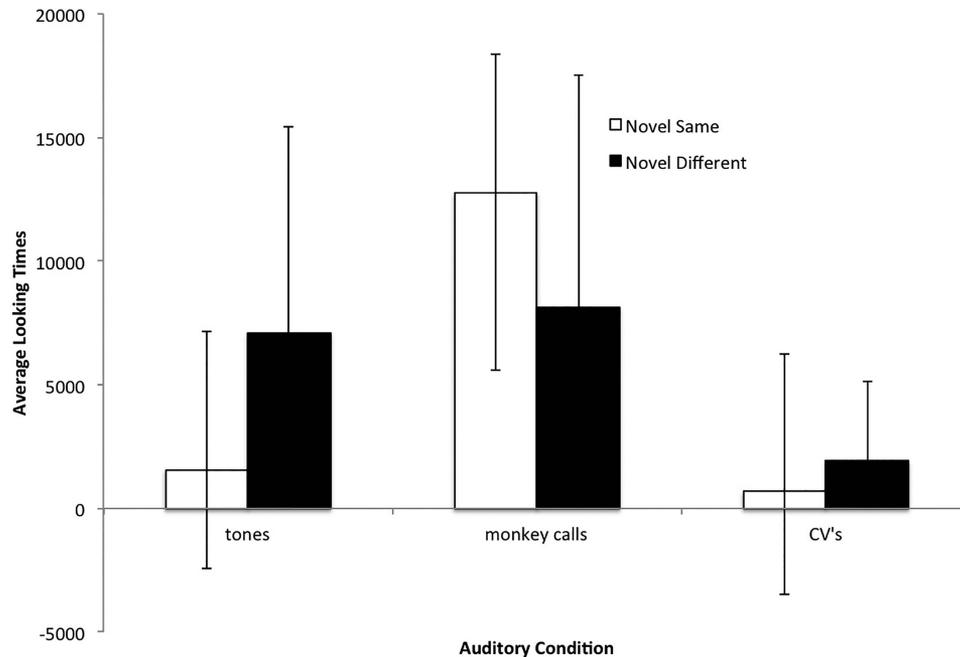


Figure 5. Average looking times and 95% confidence intervals for the novel sequence in the same pattern and the novel sequence in the different pattern for each sound type (CV's, tones, and monkey calls) from the smaller group, in which the stimuli were presented in a different order. CV = consonant-vowel.

presented last and should have profited from practice with other habituated sounds to generate quicker habituation. Rather, they never reached a level of habituation in the original group or in the smaller control group that presented monkey calls second. Grammatical structure framed as regular patterns of sounds is attended to less in these monkeys in their own “language.”

It is not clear yet the significance in human language comprehension of the respective roles of rule-based decomposition of hierarchical syntactic structures compared with an online analysis of the statistical relationship between language or sound elements. For rule-based models, the hierarchical structure of linguistic input sequence must be revealed via syntactic analysis to comprehend spoken language. In contrast, statistics-based models propose that the probabilistic relationships between adjacent or regularly patterned elements are sufficient for comprehension. In the current experiment, regular patterns between adjacent “words” did not allow for habituation to a regular pattern when monkey calls were used. The tamarins showed evidence of some degree of habituation to highly regular probabilistic relationships among human CV sounds and among tones. They were not able to habituate to the same regular probabilistic relationships in monkey calls given the same rate of exposure. Rather, the “meaning” or reaction drawn from each “word” or unit dominated their reactions.

Future tests of artificial grammar detection should consider the salience and relevance of different sounds to the organism and should test both neutral artificial constructions as well as structures using language-like units. If the goal is to determine if statistically regular pattern detection accounts for syntactic comprehension, then similar outcomes across language-like units and more neutral or artificial units should be found. In contrast, if it becomes clear

that other species of primates and possibly other animals do not use the same pattern detection ability in their own communication and they can do so with artificial stimuli, but that humans can do so across all sonic domains including their own language sounds, then humans are applying statistically based models more uniformly across language, music, and artificial sounds. Tamarin communicative sounds may not be interpreted by tamarins in this way.

References

- Abe, K., & Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nature Neuroscience*, *14*, 1067–1074. <http://dx.doi.org/10.1038/nn.2869>
- Bates, E., & Elman, J. (1996). Learning rediscovered. *Science*, *274*, 1849–1850. <http://dx.doi.org/10.1126/science.274.5294.1849>
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press, Print.
- Corballis, M. (2009). Do rats learn rules? *Animal Behaviour*, *78*, e1–e2. <http://dx.doi.org/10.1016/j.anbehav.2009.05.001>
- Ding, N., Melloni, L., Tian, X., & Poeppel, D. (2016). Rule-based and word-level statistics-based processing of language: Insights from neuroscience. *Language, Cognition and Neuroscience*. Advance online publication. <http://dx.doi.org/10.1080/23273798.2016.1215477>
- Ferguson, B., & Lew-Williams, C. (2016). Communicative signals support abstract rule learning by 7-month-old infants. *Scientific Reports*, *6*, 25434. <http://dx.doi.org/10.1038/srep25434>
- Ghitza, O. (2016). Acoustic-driven delta rhythms as prosodic markers. *Language, Cognition and Neuroscience*. Advance online publication. <http://dx.doi.org/10.1080/23273798.2016.1232419>
- Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186. [http://dx.doi.org/10.1016/S1364-6613\(00\)01467-4](http://dx.doi.org/10.1016/S1364-6613(00)01467-4)

- Gonzales, K., Gerken, L., & Gómez, R. L. (2015). Does hearing two dialects at different times help infants learn dialect-specific rules? *Cognition*, *140*, 60–71. <http://dx.doi.org/10.1016/j.cognition.2015.03.015>
- Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, B53–B64. [http://dx.doi.org/10.1016/S0010-0277\(00\)00132-3](http://dx.doi.org/10.1016/S0010-0277(00)00132-3)
- Hauser, M., Weiss, D., & Marcus, G. (2002). Rule learning by cotton-top tamarins. *Cognition*, *86*, B15–B22. [http://dx.doi.org/10.1016/S0010-0277\(02\)00139-7](http://dx.doi.org/10.1016/S0010-0277(02)00139-7)
- Hauser, M., Weiss, D., & Marcus, G. (2010). RETRACTED: Rule learning by cotton-top tamarins. *Cognition*, *117*, 106.
- Herbranson, W. T., & Shimp, C. P. (2008). “Artificial grammar learning” in pigeons: A preliminary analysis. *Learning & Behavior*, *31*, 98–106. <http://dx.doi.org/10.3758/BF03195973>
- Honing, H., ten Cate, C., Peretz, I., & Trehub, S. E. (2015). Without it no music: Cognition, biology and evolution of musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*, 20140088. <http://dx.doi.org/10.1098/rstb.2014.0088>
- Marcus, G. F., Fernandes, K. J., & Johnson, S. P. (2007). Infant rule learning facilitated by speech. *Psychological Science*, *18*, 387–391. <http://dx.doi.org/10.1111/j.1467-9280.2007.01910.x>
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80. <http://dx.doi.org/10.1126/science.283.5398.77>
- Marler, P. (2000). Origins of tones and speech: Insights from animals. In N. Wallin, B. Merker, & S. Brown (Eds.), *The origins of tones* (pp. 31–48). Cambridge, MA: MIT Press.
- Murphy, R. A., Mondragón, E., & Murphy, V. A. (2008). Rule learning by rats. *Science*, *319*, 1849–1851. <http://dx.doi.org/10.1126/science.1151564>
- Neiwirth, J. J. (2013). Chasing sounds. *Behavioural Processes*, *93*, 111–115. <http://dx.doi.org/10.1016/j.beproc.2012.11.009>
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Rabagliati, H., Senghas, A., Johnson, S., & Marcus, G. F. (2012). Infant rule learning: Advantage language, or advantage speech? *PLoS ONE*, *7*, e40517. <http://dx.doi.org/10.1371/journal.pone.0040517>
- Saffran, J. R. (2002). Constraints on statistical language learning. *Journal of Memory and Language*, *47*, 172–196. <http://dx.doi.org/10.1006/jmla.2001.2839>
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. <http://dx.doi.org/10.1126/science.274.5294.1926>
- Saffran, J., Hauser, M., Seibel, R., Kapfhamer, J., Tsao, F., & Cushman, F. (2008). Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*, *107*, 479–500. <http://dx.doi.org/10.1016/j.cognition.2007.10.010>
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52. [http://dx.doi.org/10.1016/S0010-0277\(98\)00075-4](http://dx.doi.org/10.1016/S0010-0277(98)00075-4)
- Spierings, M. J., & ten Cate, C. (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, E3977–E3984.
- ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: Natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *367*, 1984–1994. <http://dx.doi.org/10.1098/rstb.2012.0055>
- Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, *67*, 867–875. <http://dx.doi.org/10.3758/BF03193539>
- van Heijningen, C. A., Chen, J., van Laatum, I., van der Hulst, B., & ten Cate, C. (2013). Rule learning by zebra finches in an artificial grammar learning task: Which rule? *Animal Cognition*, *16*, 165–175. <http://dx.doi.org/10.1007/s10071-012-0559-x>
- Yang, C. D. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences*, *8*, 451–456. <http://dx.doi.org/10.1016/j.tics.2004.08.006>
- Zhou, H., Melloni, L., Poeppel, D., & Ding, N. (2016). Interpretations of frequency domain analyses of neural entrainment: Periodicity, fundamental frequency, and harmonics. *Frontiers in Human Neuroscience*, *10*, 274. <http://dx.doi.org/10.3389/fnhum.2016.00274>

Received August 2, 2016

Revision received January 23, 2017

Accepted January 24, 2017 ■