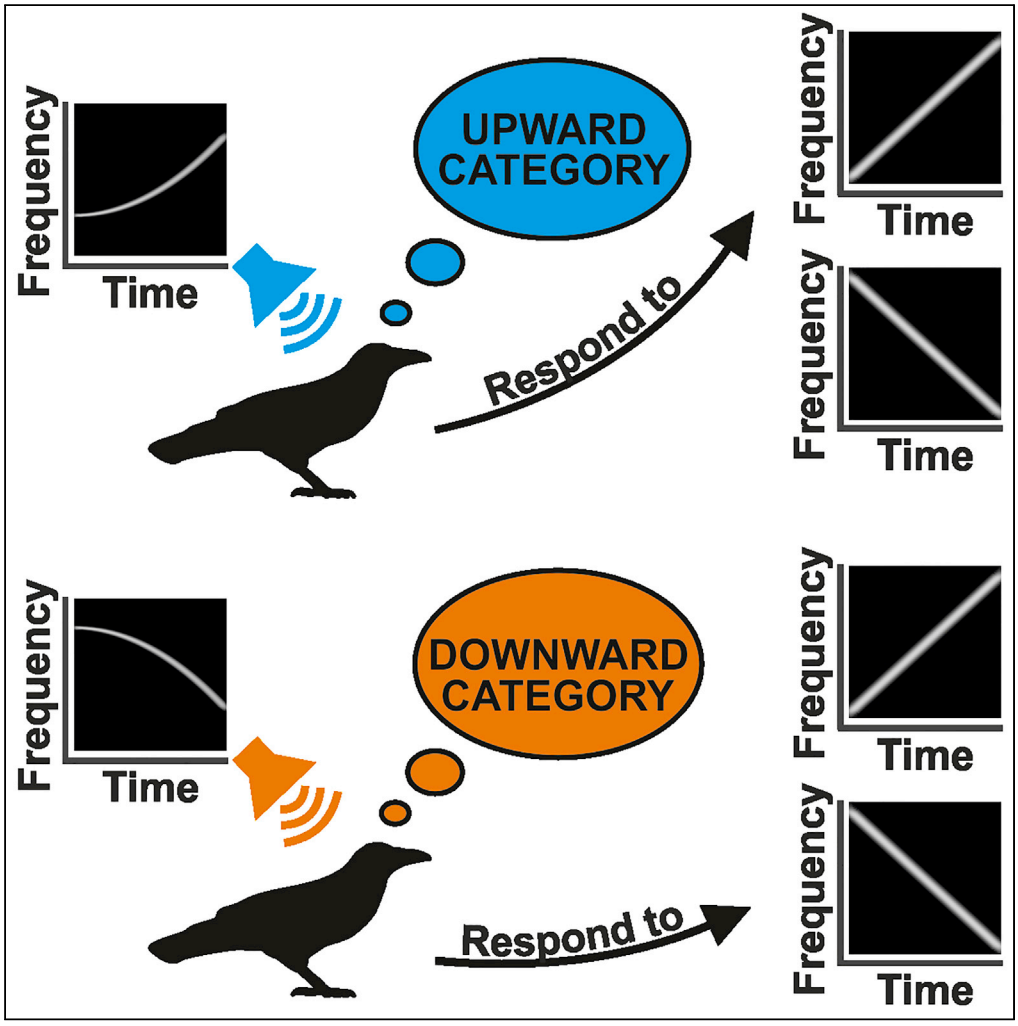
iScience

CellPress
OPEN ACCESS

Article

Categorical Auditory Working Memory in Crows



Lysann Wagener, Andreas Nieder

andreas.nieder@uni-tuebingen.de

HIGHLIGHTS

Crows performed a delayed match-to-category task with frequency modulated sounds

Crows classified novel sounds into upward or downward modulated sound categories

Crows showed sharp category boundaries and within-category generalization

Crows can actively memorize auditory perceptual categories for cognitive control

Wagener & Nieder, iScience *23*, 101737
November 20, 2020
https://doi.org/10.1016/j.isci.2020.101737

iScience

CellPress
OPEN ACCESS

Article

Categorical Auditory Working Memory in Crows

Lysann Wagener[1] and Andreas Nieder[1,2,*]

SUMMARY

The ability to group sensory data into behaviorally meaningful classes and to maintain these perceptual categories active in working memory is key to intelligent behavior. Here, we show that carrion crows, highly vocal and cognitively advanced corvid songbirds, possess categorical auditory working memory. The crows were trained in a delayed match-to-category task that required them to flexibly match remembered sounds based on the upward or downward shift of the sounds' frequency modulation. After training, the crows instantaneously classified novel sounds into the correct auditory categories. The crows showed sharp category boundaries as a function of the relative frequency interval of the modulation. In addition, the crows generalized frequency-modulated sounds within a category and correctly classified novel sounds kept in working memory irrespective of other acoustic features of the sound. This suggests that crows can form and actively memorize auditory perceptual categories in the service of cognitive control of their goal-directed behaviors.

INTRODUCTION

Categorical working memory, the ability to group sensory data into behaviorally meaningful classes and to maintain them active in working memory for a future goal, is key to intelligent behavior (Miller et al., 2018). It allows humans and animals to classify, memorize, and process sensory information efficiently. This enables humans and cognitively advanced animals to quickly adapt to new situations (Miller et al., 2003).

So far, categorical working memory in animals has primarily been demonstrated in the visual domain. In classical working memory tasks, monkeys and crows flexibly switch between remembered visual categories, such as "leftward versus rightward motion" (Zhou and Freedman, 2019), "cats versus dogs" (Freedman et al., 2001), or "same versus different" (Wallis et al., 2001; Veit and Nieder, 2013). However, whether categorical working memory is also found in the auditory domain is currently unknown.

This lack of knowledge about auditory categorical working memory is surprising because this cognitive capability is essential during goal-directed audio-vocal communication. In a telephone group call, for instance, we categorize speech signals as belonging to a specific individual and maintain this auditory category in working memory in order to match it to subsequent speech signals of the same speaker while following a conversation. Undoubtedly, also animals that rely on elaborate audio-vocal communication would benefit from this cognitive ability. Unfortunately, most animals are notoriously difficult to train on complex auditory tasks (Plakke and Romanski, 2016). Currently it is therefore rarely studied whether animals can actively maintain auditory categories in working memory (Tsunada et al., 2011).

As true vocal learners, songbirds face many challenges of acoustic communication with speaking humans (Mooney, 2009). To follow an audio-vocal communication, songbirds need to recognize communication partner's characteristics, such as sex, group membership, or identity (Wascher et al., 2015; Brecht and Nieder, 2020). In short, songbirds rely both on acute hearing and cognitive abilities to classify a multitude of raw acoustic stimuli and memorize this information across time (Nieder and Mooney, 2020). Indeed, songbirds are known to perceive sounds in a categorical way (Dooling et al., 1995; Burgering et al., 2019). In addition, they show working memory for auditory items comparable with humans (Zokoll et al., 2007; Comins and Gentner, 2010). However, whether birds can combine both capabilities to actively memorize auditory categories for future goal-directed behavior is unknown, and this capability is barely studied in animals in general. Here, we addressed this issue in carrion crows, a vocal corvid songbird

[1]Animal Physiology Unit, Institute of Neurobiology, University of Tübingen, Auf der Morgenstelle 28, 72076 Tübingen, Germany

[2]Lead Contact

*Correspondence: andreas.nieder@uni-tuebingen.de

https://doi.org/10.1016/j.isci.2020.101737

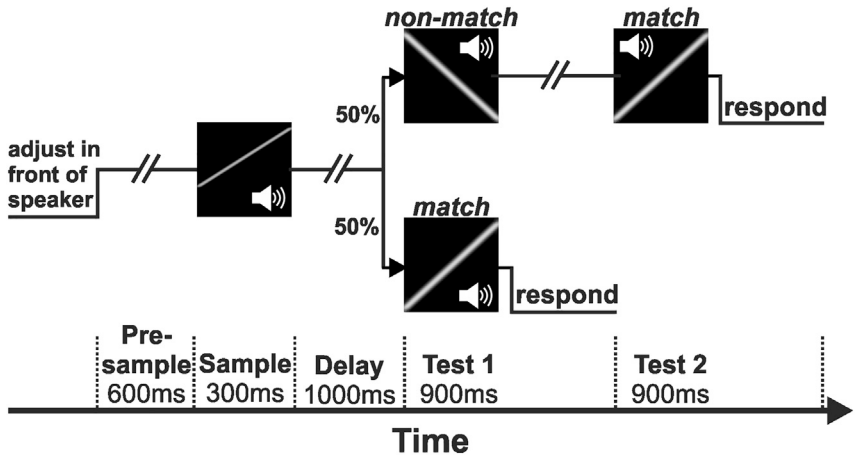


Figure 1. Task Design

The trial began when the crow adjusted its head in front of the speaker and screen (by entering an infra-red light barrier) in response to a central visual Go-cue displayed on the screen. After the crow had adjusted its head, the screen turned blank for the rest of the trial. A silent pre-sample period (600 ms) was followed by a frequency-modulated sample sound that was played for 300 ms. The sample was followed by a 1s silent delay and then by a choice (Test) sound (900 ms). *Lower trial end-sequence:* If the category (upward or downward FM) of Test1 matched that of the sample ("match" condition), the crow had to move its head and leave the infra-red light barrier to the Test1 sound within the 900 ms response time (shifted by 100 ms relative to Test-onset) to obtain a food reward. *Upper trial end-sequence:* If Test1 was a nonmatch ("non-match" condition), a match followed as Test2, which required a head movement for a reward. There were an equal number of match and nonmatch trials and they were randomly interleaved.

that can be trained on complex tasks (Nieder, 2017; Brecht et al., 2019; Nieder et al., 2020) requiring conceptual understanding and behavioral flexibility (Veit et al., 2015; Moll and Nieder, 2014; Smirnova et al., 2015; Ditz and Nieder, 2016a).

RESULTS

We trained crows on a delayed match-to-category task with sounds (Figure 1). In this task, the crows indicated whether a test sound was a categorical match to a previously presented and memorized sample sound. In each trial, the crows evaluated and maintained the direction of frequency modulation (FM) of the sample sounds in working memory to subsequently match them to the upward or downward modulated sound categories. Since individual trials presented varying sound combinations, the crows had to flexibly categorize what they heard on a trial-by-trial basis.

The crows were first trained to match six fixed FM sample stimuli ("training stimuli," three upward and three downward sweeps) to the upward or downward categories (Figures 2A and 2B). The frequency range of the upward and downward FM stimuli together covered the entire hearing range of crows (Jensen and Klokker, 2006). Once the crows reached reliable performance with these training sample stimuli, novel probe sample stimuli were occasionally inserted in the daily sessions (Figures 2C–2E), while the crows continued to discriminate the training stimuli as background task. Both crows performed 10 successive sessions with randomly interleaved training and probe stimuli.

For the training sample stimuli, crow O performed an average of 430 correct background trials per session ($\pm$52 STD, n = 10) and reached mean performance of 85.2% ($\pm$6.1% STD across sessions) (Figure 3). Crow G on average accomplished 426 correct background trials per session ($\pm$36 STD, n = 10), with a mean performance of 87.7% ($\pm$2.5% STD) (Figure 3). The average performance of both crows with the background stimuli in each daily session was significantly above the 50%-chance level (each binomial test, $p < 0.001$). Owing to the temporal succession of the matching test stimulus in the "match" versus "non-match conditions," both crows had a bias toward responding to test1, resulting in systematically higher performances during match trials (see separate data points for match and non-match performances in Figure 3). However, not only match but also all non-match performances separately were significantly above chance for both crows and all conditions (each binomial test, $p < 0.001$). The crows' mean performances for each of the six training sample stimuli was indifferent (each one-way ANOVA, $p > 0.05$).

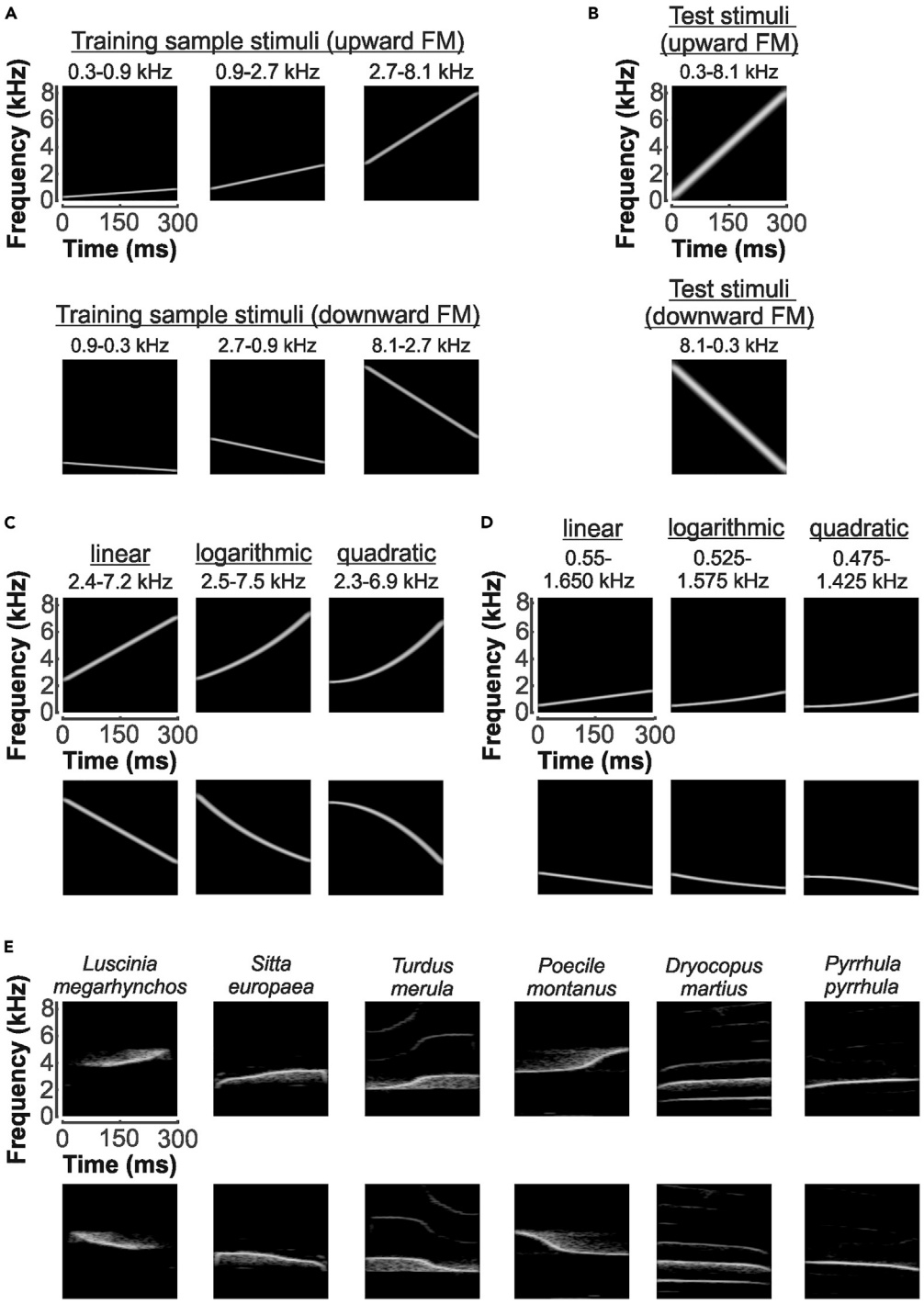


Figure 2. Auditory Stimuli Depicted as Sonagrams

(A) The familiar sample stimuli for training the crows were three upward and three downward linear FM sweeps.

(B) The same two upward and downward FM sweeps were used as test stimuli.

(C) Examples of new probe sample stimuli with frequency interval ratios of 3:1 (1.6 octaves). Linear, logarithmic, and quadratic sweeps in a high-frequency range are shown. Top row displays upward FM sweeps, bottom row shows the corresponding downward FM sweeps.

(D) Same layout as in (C), showing linear, logarithmic, and quadratic sample probe stimuli in a low-frequency range.

(E) Probe sample stimuli consisting of segments of bird vocalizations. Six representative examples of the 20 stimuli are depicted. Top row shows upward, bottom row the corresponding mirrored downward stimuli.

Next, we tested whether the crows could generalize novel FM sounds they had never heard before to the appropriate categories and thereby would demonstrate a conceptual grasp of sound categories. To that aim, we occasionally introduced novel probe sample sounds (12% of the trials) in the daily sessions with the training sounds (the remaining 88% of the trials). Four classes of novel probe sample sounds were presented: three classes of pure-tone FM sweeps with linear (where frequency changes linearly with time), logarithmic (where frequency changes logarithmically with time), and quadratic (where frequency changes quadratically with time) frequency trajectories, and frequency-modulated segments of bird vocalizations. The frequency interval ratios of the pure-tone probe sweeps were 2:1 (1 octave), 3:1 (1.6 octaves) (examples in Figures 2C and 2D), and 4:1 (2 octaves). The mean frequency interval ratio of probe bird vocalizations was 1.47:1 (around half an octave), on average (Figure 2E). The number of upward and downward-modulated probe stimuli was balanced. Because the goal was to test whether the crows could instantaneously transfer the FM categories without additional learning, we only analyzed responses to the first presentation of each unique probe stimulus.

Across all probe stimuli and classes, both crows showed a significant category transfer (each binomial test, $p < 0.001$, $n = 160$) (Figure 4). For all ten sessions together, crow O responded 80% (128/160 trials) and crow G responded 77% (123/160) correctly across all probe stimuli (which was comparable with the performance with training stimuli in crow O but significantly worse in crow G; binomial test, $p < 0.05$). To ensure that the transfer was made for each of the two categories, we analyzed the performance to upward and downward FM probes separately. Again, both crows performed well above chance level for both categories separately (each binomial test, $p < 0.01$, $n = 80$) (Figure 4). Crow O responded correctly in 81% and 79% of the trials presenting upward and downward FM probe stimuli, respectively. Crow G responded correctly in 68% and 86% of the trials presenting upward and downward FM probe stimuli, respectively. Again, not only match but also non-match performances separately were significantly above chance for both crows and all conditions (each binomial test, $p < 0.05$), except for one (downward for crow O, binomial test, $p = 0.059$).

Categorization is characterized by sharp category boundaries and within-category generalization. We first analyzed performance as a function of distance to the category boundary. The physical dimension for categorization of FM sounds into the perceptual "upward" and "downward" categories is the frequency interval ratio of the sounds. A frequency interval ratio of 1 (i.e., no change in frequency with time) demarcates the category boundary relative to which upward versus downward frequency-modulated sounds of increasing frequency interval ratio can be classified into the FM categories upward versus downward. Figure 5 depicts the crows' judgments of upward category as a function of the probes' frequency interval ratios. As expected for categorical behavior, the crows classified rising FM sounds into the upward category and falling FM sounds into the downward category, with an abrupt switch of performance at the category boundary. Performance for probe sweeps at high-frequency interval ratios (4:1, 3:1, and 2:1) (each binomial test, $p < 0.001$, $n = 30$ for ratios of 4:1 and 2:1, respectively, $n = 60$ for a ratio of 3:1). The performance of crow O was 93%, 75%, and 90% for ratios of 4, 3, and 2, respectively. The performance of crow G was 80%, 85%, and 87% for ratios of 4, 3, and 2, respectively. As expected, categorization with probe bird vocalizations that had the lowest frequency interval ratio of all probe sounds near the category boundary became increasingly more difficult for the crows. Crow O correctly categorized the probe bird vocalization sounds (70%; binomial test, $p < 0.01$, $n = 40$), whereas crow G showed a tendency but did not reach significance (55%; binomial test, $p = 0.32$, $n = 40$). Overall, however, the crows categorized novel sounds correctly into the appropriate categories, with categorization performance suffering close to the category boundary.

Next, we investigated within-category generalization performance. Within-category generalization predicts that performance is independent from the acoustic details of the FM sound, such as the modulation trajectory and the frequency composition of the sounds. To that aim, we separately analyzed and compared performance to the four probe stimulus classes (linear, logarithmic, quadratic pure-tone FM sweeps, and bird vocalization segments). Both crows showed high performance to all probes containing FM sweeps of different trajectories. (linear: crow O 85%, crow G 85%; logarithmic: crow O 85%, crow G 83%; quadratic: crow O 80%, crow G 85%) (each binomial test, $p < 0.001$, $n = 40$) (Figure 6). As mentioned above, the bird vocalization probes that exhibited only mild frequency modulation were close to the category boundary and thus more difficult for the crows. To summarize, for probe sounds with distinct frequency modulation, the crows categorized performance was independent from the type of modulation trajectory.

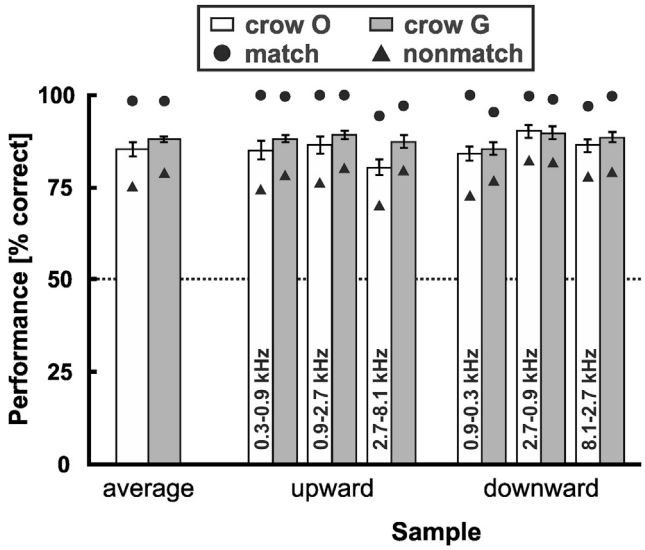


Figure 3. Performance to Familiar Training Stimuli

Both crows responded significantly above chance (dashed horizontal line at 50% performance) to upward and downward FM samples. Columns represent mean performance values averaged across match and non-match trials (error bars: standard error of the mean), circle and triangle symbols reflect mean performance for match and non-match trials separately.

In addition, we investigated whether the frequency range of the 120 pure-tone probe stimuli (linear, logarithmic, and quadratic sweeps) had an influence on behavior. Half of these stimuli had a frequency between 0.3 and 2.7 kHz and were therefore assigned to the group of "low-frequency" stimuli. The other half had a frequency between 0.9 and 8.1 kHz and were grouped as "high-frequency" stimuli. Stimuli including frequencies in the overlapping range of 0.9–2.7 kHz never contained both frequencies lower than 0.9 kHz and higher than 2.7 kHz. The crows performed well above chance regardless of the frequency range of the sample stimuli (each binomial test, $p < 0.001$, $n = 60$) (Figure 6). Crow O responded correctly in 87% and 80% of low frequency and high frequency trials, respectively. Crow G responded correctly in 92% and 77% of low frequency and high frequency trials, respectively. Thus, the crows showed robust within-category generalization irrespective of the frequency range of the probe sounds.

DISCUSSION

Our data show that crows possess categorical auditory working memory. They are able to maintain the FM categories upward and downward in working memory to master an auditory delayed match-to-category task. As a sign of categorical generalization and transfer, the crows instantaneously and without further training matched the remembered novel sample sounds correctly to the upward and downward FM categories, irrespective of other sound parameters. The crows' behavior showed the diagnostic characteristics of categories, namely, sharp category boundaries and within-category generalization: the crows categorically classified the continuous direction of FM into upward and downward while ignoring other sound parameters (such as spectral composition, frequency intervals, or modulation trajectory of the novel sample sounds) within one FM sound category. This suggests that the crows only memorized the direction of the FM, not the other varying sound parameters, when categorizing sounds from working memory.

Auditory Categorization in Birds

Birds have also been shown to discriminate and classify complex sounds. Vocal learners, in particular, rely on acute audition and are known to perceive sounds in a categorical way (Dooling et al., 1995; Burgering et al., 2019). Even pigeons, non-songbirds with an unlearned vocal repertoire, are able to make same/different discriminations across a wide variety of auditory stimuli (Murphy and Cook, 2008; Cook and Brooks, 2009; Cook et al., 2016) and can learn to discriminate among music-derived acoustic elements and sequences (Brooks and Cook, 2010; Hagmann and Cook, 2010; Brooks and Cook, 2010; Cook, 2017). However, previous experiments did not require the birds to flexibly switch between auditory

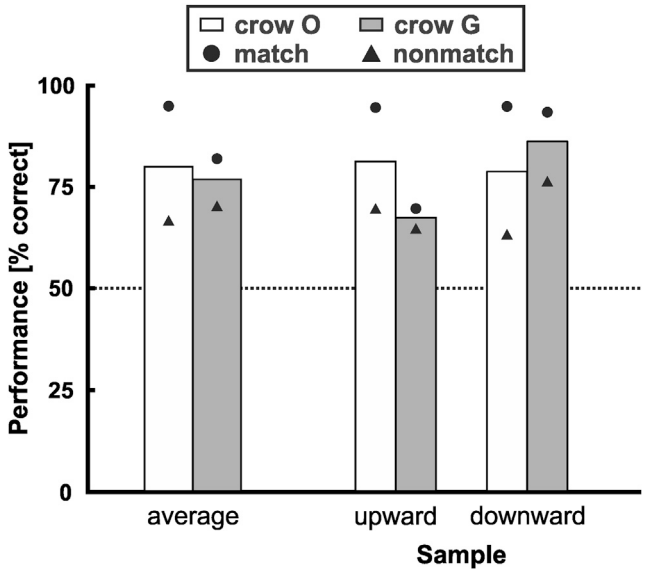


Figure 4. Overall Performance to Novel Probe Stimuli

Both crows responded significantly above chance to upward and downward FM probe samples. Columns represent mean performance values averaged across match and non-match trials (error bars: standard error of the mean), circle and triangle symbols reflect mean performance for match and non-match trials separately.

categories or remember auditory categories in working memory. In these studies, the birds were typically tested in Go/NoGo or forced choice tasks without a delay period. Both temporal and spectral changes in the sounds could be exploited.

Birds are known to categorize complex sounds, such as human speech sounds, based on temporal differences. For instance, budgerigars place vowels /i/, /a/, /e/, and /u/ in phonetically appropriate categories in spite of variation in who is talking and their gender (Dooling and Brown, 1990). When working with synthetic phoneme continua of speech sounds, budgerigars exhibit perceptual phonemic boundaries near the human boundaries for /ba/-/pa/, /da/-/ta/, /ga/-/ka/, /ra/-/la/, and /ba/-/wa/ (Dooling et al., 1995; Dent et al., 1997). Similar perception of speech sound categories has also been shown in quails and zebra finches (Burgering et al., 2019; Kleunder et al., 1987; Ohms et al., 2010). Because the phoneme boundaries rely on temporal differences (or "voice onset time" between the vowel and the consonant), these data suggest that not only sound frequency but also sound timing plays an important role in birds' capability to categorize sounds.

Besides temporal factors, also the spectral composition of sounds can be exploited by birds. In a series of experiments, several songbird species (primarily European starlings) have been shown to perceive pitch relations in a simple tonal melody (Hulse and Cynx, 1985). In particular, songbirds can classify rising as opposed to falling pitch patterns. However, these songbirds preferentially discriminated tonal patterns according to the absolute frequency of the individual element tones in the patterns; they failed to transfer discrimination to a novel frequency range when the training frequency range was shifted. Only when the experimental conditions severely constrained the use of pattern element cues did the songbirds use pitch relations as a secondary strategy (Hulse and Cynx, 1986; Hulse et al., 1984; Braaten et al., 1990). Data like these lead to the conclusion that birds, unlike humans, cannot generalize relative pitch discrimination to new frequencies, thus lacking a conceptual grasp of frequency modulation in complex sounds. However, our data suggest that corvid songbirds can indeed form a conceptual understanding of upward and downward frequency modulation, irrespective of frequency composition.

Auditory Working Memory in Birds

Auditory working memory capabilities have only rarely been studied in birds, mainly because it is difficult to train birds—and nonhuman animals in general—to perform auditory working memory tasks that are similar

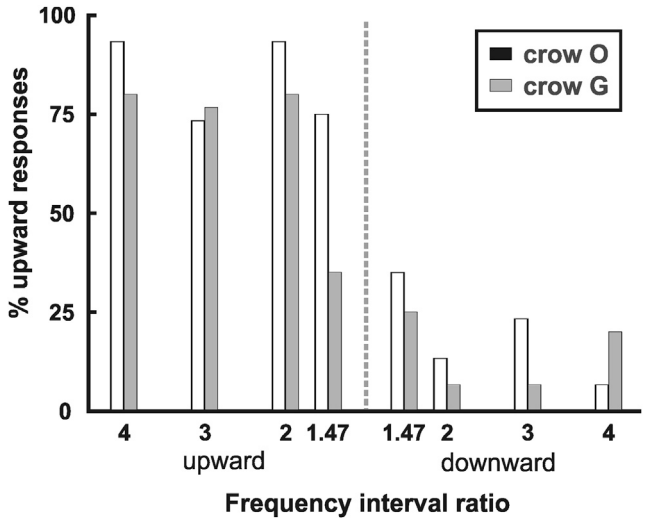


Figure 5. Performance Relative to Category Boundary

Categorization performance to probe stimuli of different frequency interval ratios of upward and downward FM sounds. Performance is depicted as percent correct classification as "upward" category. Vertical dashed line indicates the category boundary at a frequency interval ratio of 1.

to those used in the study of visual memory (Plakke and Romanski, 2016). Nonetheless, a few studies show that European starlings exhibit auditory working memory and show interesting similarities and differences when compared with humans (Zokoll et al., 2007, 2008a, 2008b; Comins and Gentner, 2010). For example, the classical finding of a decay of working memory with increasing delay times in humans and other animals could be reproduced in starlings (Zokoll et al., 2008a, 2008b). In contrast to humans, however, starlings benefited from repeated presentations of sample sounds. Our study adds to these insights by showing that songbirds maintain not only specific sounds in working memory but also overarching auditory categories. Overall, songbirds are therefore valuable models for investigating not only mechanisms of auditory signal processing but also cognitive control functions in the auditory domain.

Categorization of Bird Vocalizations

In contrast to novel pure-tone FM sweeps, novel segments of frequency-modulated bird vocalizations were more difficult to categorize for the crows. One crow reached significant categorization (albeit with less precision than with the pure-tone probes), whereas the other crow showed a tendency but failed significance. Most likely, this difficulty was due to the vocalization segments having the lowest frequency interval ratio of all probe sounds, a ratio that was closest to the category boundary. In addition, the vocalizations were acoustically more complex and richer. Some of them contained broadband noise that potentially could have masked the FMs and additional harmonics that might have distracted the crows. Overall, however, these data suggest that corvids can categorize and remember animal sounds in order to adapt their behavior.

The capability to memorize sound categories may also have adaptive advantages in a world in which objects and events are characterized by multi-modal signals. The semantic grouping of a multitude of unique stimuli into uni-modal categories facilitates the association with stimuli from other sensory modalities that characterize the same members of a class. For instance, social songbirds need to group conspecifics into different categories based on sex, relatedness, or group membership in order to adjust their behavioral responses. Crows recognize group members by identity congruence between visual presentation of a group member and the subsequent playback of a contact call (Kondo et al., 2012). Because corvids can recognize individuals by sound (Wascher et al., 2015) or sight alone (Kondo et al., 2010), the most parsimonious explanation is that they first categorize acoustic and visual stimuli as belonging to an individual and later associate the auditory and visual categories for cross-modal audiovisual recognition of group members. The brain of crows is able to associate stimuli across modality and time (Moll and Nieder, 2015, 2017). However, whether this extends also to more cognitive cross-modal categories remains to be explored.

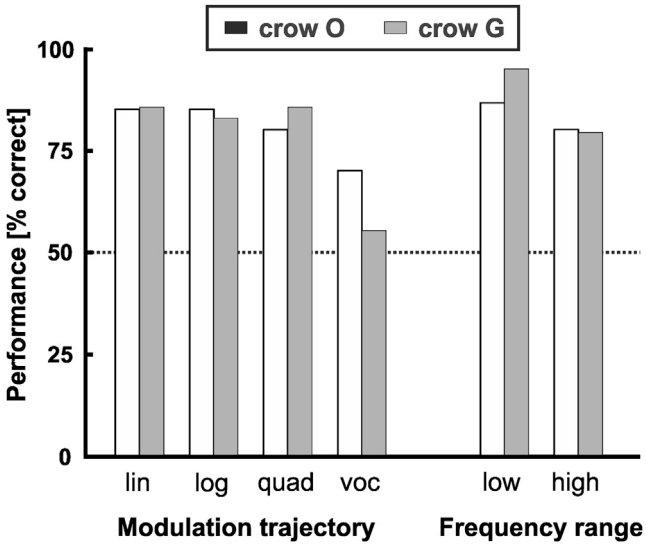


Figure 6. Performance to Probe Stimuli as a Function of Modulation Trajectory and Frequency Range
Chance level is again 50% performance.

Categorization of Pure Auditory Frequency Modulation in Mammals

The categorical discrimination of sounds based on pure frequency modulation has been demonstrated convincingly in a mammal, the Mongolian gerbil (Wetzel et al., 1998; Ohl et al., 2001). In this positive-reinforcement Go/NoGo task, the effects of conditioned fear (CS+) based on FM categories were tested. The gerbils had to change compartments in a shuttle box during ascending FMs (CS+) presentation to avoid foot shock. The gerbils were able to discriminate FM tones by modulation direction and, after familiarization with a number of different FM pairs, transferred the ascending-descending concept to stimuli not heard before (Wetzel et al., 1998). A similar conditioning approach was used in categorization studies with ferrets (Yin et al, 2016, 2020); in one study, individual ferrets were trained to discriminate downward sequences (the target sequence) from upward sequences (the reference sequence), or vice versa (Yin et al., 2010). In both approaches, gerbils and ferrets thus discriminated a fixed FM category stored in long-term memory from deviating sounds.

Although these experiments clearly show perceptual categorization of FMs in gerbils and ferrets, they required the animals neither to flexibly switch between different auditory categories nor to maintain the switching categories in auditory working memory. To address both cognitive aspects, we therefore trained crows on a delayed match-to-category task. This task not only tested the formation of one FM category against other sounds but probed the conceptual flexibility of the crows to switch between rewarded and unrewarded FM categories on a trial-by-trial basis. In addition, the crows could not have succeeded without a working memory for the auditory categories.

Categorical Auditory Working Memory in Monkeys

Categorical auditory perception and working memory have been reported in macaque monkeys. Using a delayed match-to-sample protocol, monkeys were trained to report by an eye movement whether two consecutive human-speech sounds (“dad” versus “bad”) or a series of morphed versions of these sounds belonged to the same or different category (Tsunada et al., 2011). The behavioral data showed that monkeys perceived these morphed speech sounds categorically; despite the gradual variation of the acoustic stimulus, the monkeys reliably assigned the morphs to one of the two categories and exhibited a sharp transition boundary between morphed sounds being perceived as dad rather than bad.

Whether the monkeys could also categorize novel morph sounds or other types of speech sounds as a sign of abstract categorization was not tested in this study. We tested this in the current study and found that the crows instantaneously categorized the remembered novel sample sounds correctly to the upward and

downward FM categories, irrespective of other sound parameters. Crows can transfer the semantic grouping criteria they learned to novel and acoustically distinct sounds.

It is worth mentioning that the auditory working memory capacity of monkeys seems to be surprisingly limited and prone to interference. When rhesus monkeys were tested in an auditory delayed match-to-sample task equivalent to the task structure of the current study in which either the first (match condition) or the second test stimulus (nonmatch condition) could be a match and required a response, marked performance differences between the two conditions surfaced. Performance was accurate whenever a match followed the sample directly, but it fell precipitously if (one or two) nonmatch stimuli intervened between sample and match. This drop in accuracy was found to result from an "overwriting" effect, i.e., a retroactive interference from the intervening nonmatch stimulus that was far greater than that observed previously in delayed match-to-sample tasks with visual stimuli. The authors concluded that the monkeys' performance depended on the retention of stimulus traces in the passive form of short-term memory rather than on active working memory (Scott et al., 2012, 2013).

Our data from crows only allow an evaluation of this issue for zero (match condition) or one interfering stimulus (nonmatch condition). The data plotted in Figures 3 and 4 show a similar tendency, namely, a decline in accuracy in the nonmatch condition. Notably, crow G showed only a mild decline in the nonmatch condition when tested with novel probe stimuli (Figure 4). It is also worth mentioning that part (or all) of this performance decline may be due to the crows' bias to respond rather quicker (match condition) to receive a reward earlier. In addition, the performance and response pattern of crows for match and nonmatch conditions is comparable with those we see for visual categorization in delayed match-to-sample tasks (Ditz and Nieder, 2016a, 2016b, 2020; Wagener et al., 2018). Overall, the data suggest that the crows possess active working memory capacities also for auditory stimuli.

Limitations of the Study

This study explored the crows' category generalization capabilities to a limited set of probe stimuli and found that the crows had more difficulty categorizing FM segments of bird vocalizations. One explanation for this finding is that vocalizations showed the smallest frequency interval ratio of all probe stimuli. However, compared with the pure tone training FM sweeps, vocalizations also showed additional harmonics. To demonstrate that crows can generalize FM categories to acoustically richer sounds, the application of multi-harmonic FM sweeps as training and probe stimuli would be helpful. In addition, and to further differentiate active working memory from potential passive short-term memory, the crows' performance when confronted with more than one distractor and for longer delays would be informative. Resistance against distraction over longer delay periods would corroborate the notion of auditory working memory in crows as it is regularly seen in the visual domain.

Resource Availability

Lead Contact

Andreas Nieder, andreas.nieder@uni-tuebingen.de.

Materials Availability

This study did not generate new unique Materials.

Data and Code Availability

Original data have been deposited to Mendeley: https://doi.org/10.17632/38x8ktrx7y.1.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.101737.

ACKNOWLEDGMENTS

This work was supported by a DFG grant NI 618/11-1 to A.N.

AUTHOR CONTRIBUTIONS

A.N. and L.W. designed the study, interpreted the data and wrote the manuscript. L.W. performed experiments and analyzed the data. All authors gave final approval for publication.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 4, 2020
Revised: September 10, 2020
Accepted: October 23, 2020
Published: November 20, 2020

REFERENCES

Braaten, R.F., Hulse, S.H., and Page, S.C. (1990). Discrimination and classification of rising and nonrising pitch patterns by the European starling (*Sturnus vulgaris*). Anim. Learn. Behav. *18*, 352–364.

Brecht, K.F., Hage, S.R., Gavrilov, N., and Nieder, A. (2019). Volitional control of vocalizations in corvid songbirds. PLoS Biol. *17*, e3000375.

Brecht, K.F., and Nieder, A. (2020). Parting self from others: individual and self-recognition in birds. Neurosci. Biobehav. Rev. *116*, 99–108.

Brooks, D.I., and Cook, R.G. (2010). Chord discrimination by pigeons. Music Percept. *27*, 183–196.

Burgering, M.A., Vroomen, J., and ten Cate, C. (2019). Zebra finches (*Taeniopygia guttata*) can categorize vowel-like sounds on both the fundamental frequency ("pitch") and spectral envelope. J. Comp. Psychol. *133*, 106–117.

Comins, J.A., and Gentner, T.Q. (2010). Working memory for patterned sequences of auditory objects in a songbird. Cognition *117*, 38–53.

Cook, R.G. (2017). Music perception in a comparative context: relational chord perception by pigeons. Comp. Cogn. Behav. Rev. *12*, 45–56.

Cook, R.G., and Brooks, D.I. (2009). Generalized auditory same-different discrimination by pigeons. J. Exp. Psychol. Anim. Behav. Process. *35*, 108–115.

Cook, R.G., Qadri, M.A.J., and Oliveira, R. (2016). Detection and discrimination of complex sounds by pigeons (*Columba livia*). Behav. Process. *123*, 114–124.

Dent, M.L., Brittan-Powell, E.F., Dooling, R.J., and Pierce, A. (1997). Perception of synthetic/ba/-/wa/ speech continuum by budgerigars (*Melopsittacus undulatus*). J. Acoust. Soc. Am. *102*, 1891–1897.

Ditz, H.M., and Nieder, A. (2016a). Numerosity representations in crows obey the Weber-Fechner law. Proc. Biol. Sci. *283*, 20160083.

Ditz, H.M., and Nieder, A. (2016b). Sensory and working memory representations of small and large numerosities in the crow endbrain. J. Neurosci. *36*, 12044–12052.

Ditz, H.M., and Nieder, A. (2020). Format-dependent and format-independent representation of sequential and simultaneous numerosity in the crow endbrain. Nat. Commun. *11*, 686.

Dooling, R.J., and Brown, S.D. (1990). Speech perception by budgerigars (*Melopsittacus undulatus*): spoken vowels. Percept. Psychophys. *47*, 568–574.

Dooling, R.J., Best, C.T., and Brown, S.D. (1995). Discrimination of synthetic full-formant and sinewave/ra-la/continua by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Taeniopygia guttata*). J. Acoust. Soc. Am. *97*, 1839–1846.

Freedman, D.J., Riesenhuber, M., Poggio, T., and Miller, E.K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. Science *291*, 312–316.

Hagmann, C.E., and Cook, R.G. (2010). Testing meter, rhythm, and tempo discriminations in pigeons. Behav. Process. *85*, 99–110.

Hulse, S.H., and Cynx, J. (1985). Relative pitch perception is constrained by absolute pitch in songbirds (Mimus, Molothrus, and Sturnus). J. Comp. Psychol. *99*, 176–196.

Hulse, S.H., and Cynx, J. (1986). Interval and contour in serial pitch perception by a passerine bird, the European starling (*Stumus vulgaris*). J. Comp. Psychol. *100*, 215–228.

Hulse, S.H., Cynx, J., and Humpal, J. (1984). Absolute and serial pitch discrimination in serial perception by birds. J. Exp. Psychol. Gen. *113*, 38–54.

Jensen, K.K., and Klokker, S. (2006). Hearing sensitivity and critical ratios of hooded crows (*Corvus corone cornix*). J. Acoust. Soc. Am. *119*, 1269–1276.

Kleunder, K.R., Diehl, R.L., and Killeen, P.R. (1987). Japanese quail can learn phonetic categories. Science *237*, 1195–1197.

Kondo, N., Izawa, E.I., and Watanabe, S. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): contact call signature and discrimination. Behaviour *147*, 1051–1072.

Kondo, N., Izawa, E.I., and Watanabe, S. (2012). Crows cross-modally recognize group members but not non-group members. Proc. Biol. Sci. *279*, 1937–1942.

Miller, E.K., Lundqvist, M., and Bastos, A.M. (2018). Working memory 2.0. Neuron *100*, 463–475.

Miller, E.K., Nieder, A., Freedman, D.J., and Wallis, J.D. (2003). Neural correlates of categories and concepts. Curr. Opin. Neurobiol. *13*, 198–203.

Moll, F.W., and Nieder, A. (2014). The long and the short of it: rule-based relative length discrimination in carrion crows, *Corvus corone*. Behav. Process. *107*, 142–149.

Moll, F.W., and Nieder, A. (2015). Cross-modal associative mnemonic signals in crow endbrain neurons. Curr. Biol. *25*, 2196–2201.

Moll, F.W., and Nieder, A. (2017). Modality-invariant audio-visual association coding in crow endbrain neurons. Neurobiol. Learn. Mem. *137*, 65–76.

Mooney, R. (2009). Neurobiology of song learning. Curr. Opin. Neurobiol. *19*, 654–660.

Murphy, M.S., and Cook, R.G. (2008). Absolute and relational control of a sequential auditory discrimination by pigeons (*Columba livia*). Behav. Process. *77*, 210–222.

Nieder, A. (2017). Inside the corvid brain—probing the physiology of cognition in crows. Curr. Opin. Behav. Sci. *16*, 8–14.

Nieder, A., and Mooney, R. (2020). The neurobiology of innate, volitional and learned vocalizations in mammals and birds. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190054.

Nieder, A., Wagener, L., and Rinnert, P. (2020). A neural correlate of sensory consciousness in a corvid bird. Science *369*, 1626–1629.

Ohl, F.W., Scheich, H., and Freeman, W.J. (2001). Change in pattern of ongoing cortical activity with auditory category learning. Nature *412*, 733–736.

Ohms, V.R., Gill, A., Van Heijningen, C.A., Beckers, G.J., and ten Cate, C. (2010). Zebra finches exhibit speaker-independent phonetic perception of human speech. Proc. Biol. Sci. *277*, 1003–1009.

Plakke, B., and Romanski, L.M. (2016). Neural circuits in auditory and audiovisual memory. Brain Res. *1640*, 278–288.

Scott, B.H., Mishkin, M., and Yin, P. (2012). Monkeys have a limited form of short-term memory in audition. Proc. Natl. Acad. Sci. U S A *109*, 12237–12241.

Scott, B.H., Mishkin, M., and Yin, P. (2013). Effect of acoustic similarity on short-term auditory memory in the monkey. Hear. Res. *298*, 36–48.

Smirnova, A., Zorina, Z., Obozova, T., and Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. Curr. Biol. *25*, 256–260.

Tsunada, J., Lee, J.H., and Cohen, Y.E. (2011). Representation of speech categories in the primate auditory cortex. J. Neurophysiol. *105*, 2634–2646.

Veit, L., and Nieder, A. (2013). Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. Nat. Commun. *4*, 2878.

Veit, L., Pidpruzhnykova, G., and Nieder, A. (2015). Associative learning rapidly establishes neuronal representations of upcoming behavioral choices in crows. Proc. Natl. Acad. Sci. U S A *112*, 15 208–215 213.

Wagener, L., Loconsole, M., Ditz, H.M., and Nieder, A. (2018). Neurons in the endbrain of numerically naive crows spontaneously encode visual numerosity. Curr. Biol. *28*, 1090–1094.e4.

Wallis, J.D., Anderson, K.C., and Miller, E.K. (2001). Single neurons in prefrontal cortex encode abstract rules. Nature *411*, 953–956.

Wascher, C.A., Hillemann, F., Canestrari, D., and Baglione, V. (2015). Carrion crows learn to discriminate between calls of reliable and unreliable conspecifics. Anim. Cogn. *18*, 1181–1185.

Wetzel, W., Wagner, T., Ohl, F.W., and Scheich, H. (1998). Categorical discrimination of direction in frequency-modulated tones by Mongolian gerbils. Behav. Brain Res. *91*, 29–39.

Yin, P., Fritz, J.B., and Shamma, S.A. (2010). Do ferrets perceive relative pitch? J. Acoust. Soc. Am. *127*, 1673–1680.

Yin, P., Shamma, S.A., and Fritz, J.B. (2016). Relative salience of spectral and temporal features in auditory long-term memory. J. Acoust. Soc. Am. *140*, 4046.

Yin, P., Strait, D.L., Radtke-Schuller, S., Fritz, J.B., and Shamma, S.A. (2020). Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. Curr. Biol. *30*, 1649–1663.e5.

Zhou, Y., and Freedman, D.J. (2019). Posterior parietal cortex plays a causal role in perceptual and categorical decisions. Science *365*, 180–185.

Zokoll, M.A., Klump, G.M., and Langemann, U. (2007). Auditory short-term memory persistence for tonal signals in a songbird. J. Acoust. Soc. Am. *121*, 2842–2851.

Zokoll, M.A., Klump, G.M., and Langemann, U. (2008a). Auditory memory for temporal characteristics of sound. J. Comp. Physiol. A *194*, 457–467.

Zokoll, M.A., Naue, N., Herrmann, C.S., and Langemann, U. (2008b). Auditory memory: a comparison between humans and starlings. Brain Res. *1220*, 33–46.

iScience, Volume *23*

Supplemental Information

Categorical Auditory Working Memory in Crows

Lysann Wagener and Andreas Nieder

Supplemental Information

Categorical auditory working memory in crows

Lysann Wagener and Andreas Nieder

TRANSPARENT METHODS

Subjects

Two 3 years old male carrion crows (*Corvus corone*) were used in this study. The crows were housed in social groups in indoor aviaries. During the training and testing period, the crows were on a controlled feeding protocol. Body weight was measured daily. Food was given as reward during the sessions. Water was *ad libitum* available in the aviary and during the experiments. All procedures were carried out according to the guidelines for animal experimentation and approved by the responsible national authorities, the Regierungspräsidium Tübingen, Germany.

Experimental setup

The birds were placed on a perch in front of a touchscreen monitor (3M Microtouch, 15“, 60 Hz refresh rate) in a darkened operant conditioning chamber (length 1 m, width 0.76 m, height 1 m). One speaker (VISATON B 200 – 6 Ohm) was used to play back the auditory stimuli. The speaker was located 0.6 m in front of the bird and behind the computer monitor. The behavior was controlled by the CORTEX system (National Institute of Mental Health, Maryland, USA) which also stored the behavioral data. An automated feeder delivered either mealworms (*Tenebrio molitor* larvae) or bird seed pellets upon correctly completed trials. An infrared light barrier was installed above the birds' head to which a reflector foil was attached. The crow had to keep its head still within the beam of the light barrier and thereby in front of the touchscreen throughout a trial.

Behavioral task

The crows were trained on a delayed match-to-category task in which they discriminated the direction of upward and downward frequency modulated (FM) sounds (**Fig. 1**). A crow started a trial by positioning its head in front of the monitor whenever a go-stimulus (small white cross) was shown on the screen. Head position was monitored by an infra-red light barrier, and the crows had to maintain the head still throughout the trial. Premature head movements terminated the trial and it was discarded. When the head was in the correct position in front of the monitor, the crows received auditory feedback and the go-stimulus on the screen turned into a white circle for 60 ms. For the further course of the trial the monitor remained black. After a 600 ms silent pre-sample phase, the auditory FM-modulated sample stimulus (300 ms duration) was played. This was followed by a 1000 ms silent delay period during which the crow had to memorize the direction of the frequency modulation (upward or downward) of the sample. In the following test phase, the crow had to match the direction of the FM in the sample to the test stimulus with the same FM direction (i.e. upward to upward FM, and downward to downward FM). If the direction of the FM matched, the crow had to respond by quickly moving its head out of the light barrier to receive a reward.

In 50% of the trials, the first test stimulus (test1) was the matching stimulus ('match condition'). In the other 50% of the trials, the test1-stimlus was a 'non-match' with a FM in the opposite direction of the sample's FM direction ('nonmatch condition'). In this case, the bird had to refrain from responding and wait with a response until the second test stimulus was played which was always a match. Both the test1- and the test2-periods were 900 ms in duration, with the 300 ms test1- and test2-stimuli played right at the beginning of the test-periods (so that the remaining 600 ms of the test-periods were silent). The response interval was shifted by 100 ms due to the inevitably reaction latency relative to physical stimulus onset. Responses to the 'nonmatch stimulus' and no response to either of the two test stimuli were considered as error and also not rewarded. Match and non-match conditions were balanced and pseudo-randomly presented. The crows were first trained with well-known training stimuli. Once the crows reached high performance, we tested if they were able to transfer the upward and downward FM categories to novel stimuli that were occasionally presented among the ongoing discrimination of the training sample stimuli.

Stimuli

A total of 168 auditory frequency modulated stimuli were used in this study. All stimuli had a duration of 300 ms and a 10 ms linear amplitude ramp at the beginning and the end.

Training stimuli. The crows were trained with a fixed set of 6 FM sample stimuli (3 upward and 3 downward sweeps). These ***training sample stimuli*** consisted of linearly rising or falling FM pure tones (**Fig. 2A**). The frequency range of the three upward training sample stimuli were 0.3-0.9 kHz, 0.9-2.7 kHz and 2.7-8.1 kHz. The identical frequency range of the three downward training sample stimuli was 0.9-0.3 kHz, 2.7-0.9 kHz and 8.1-2.7 kHz. Thus, each training sample stimulus had a bandwidth of 1.6 octaves. Each of these sample stimuli had to be matched to its corresponding matching test stimulus. A linearly FM-modulated sweep from 0.3-8.1 kHz was the match for upward FM stimuli, whereas a linear downward sweep from 8.1-0.3 kHz served as a match for downward FM stimuli (**Fig. 2B**).

Probe sample stimuli. Once the crows reliably discriminated and categorized the training stimuli, we tested their ability to transfer the upward and downward FM categories to novel sample sounds (probe stimuli). We tested a total of 80 probe stimulus pairs (each with upward and downward FM modulation) which the crows had never encountered before. Only responses to the first presentation of each unique probe stimulus – before the crows could learn a 'correct' response to these new stimuli - were analyzed. The test-stimuli remained the same as in the training trials.

The probe stimuli were grouped into four classes of FM sweeps: linear, logarithmic and quadratic FM modulation of pure tones, and FM-modulated bird vocalizations. Each of the four classes consisted of 40 unique stimuli (20 upward and 20 downward sweeps). All pure-tone sweeps (including the training, test and probe stimuli) were generated using a custom written MATLAB code. The sounds were saved as wav-files at a sampling frequency of 44.1 kHz.

$$\text{Linear: } f_i(300ms) = f_0 + \beta t, \text{ where } \beta = (f_1 - f_0)/t_1$$

$$\text{Logarithmic: } f_i(300ms) = f_0 * \beta^t, \text{ where } \beta = \left(\frac{f_1}{f_0}\right)^{\frac{1}{t_1}}$$

$$\text{Quadratic: } f_i(300ms) = f_0 + \beta t^2, \text{ where } \beta = (f_1 - f_0)/{t_1}^2$$

The pure-tone probe stimuli differed in frequency-modulation range and frequency content. The frequency-modulation ranges was quantified by the frequency interval ratio, which is the maximum frequency contained in the FM sound divided by the minimum frequency ($f_{max} : f_{min}$). The probe FM sweeps had frequency interval rations of 2:1 (1 octave), 3:1 (1.6 octaves; **Fig. 2C**) and 4:1 (2 octaves).

The frequency content was roughly divided into 'low' and 'high' frequencies. The 'low frequency' probe stimuli covered frequencies between 0.3-2.7 kHz (examples shown in **Fig. 2D**), whereas the 'high frequency' stimuli covered 0.9-8.1 kHz. Stimuli including frequencies in the overlapping range of 0.9 to 2.7 kHz were never both, lower than 0.9 kHz and higher than 2.7 kHz at once. Likewise, none of the stimuli laid exclusively within the overlap, so that each stimulus could be related to 'low' or 'high' based on whether it reached into the range of 0.3-0.9 kHz or 2.7-8.1 kHz, respectively.

The bird vocalization probe stimuli were excerpts of bird vocalizations (for example, *Parus major*, *Sturnus vulgaris*, *Buteo buteo*, *Alcedo atthis*) (downloaded from http://www.xeno-canto.org/) which have been recorded at 16-bit resolution and almost all a sampling rate of 44.1 kHz (except for two at 48 kHz and one at 16 kHz). These were further modified using Adobe Audition 3.0 and Audacity 1.0.0. From all vocalizations, a 300 ms segment covering a monotonic frequency change was extracted. The amplitude of the signal was equalized to the pure-tone stimuli and 10 ms ramps were added. Each vocalization probe stimulus was used

with its original FM-sweep direction (8/20 upward, 12/20 downward) for one FM category, and as a temporally inverted version for the other FM category. The average frequency interval ratio of the vocalization probe stimuli was 1.47:1 (± 0.25 STD).

Transfer to novel FM stimuli was tested during 10 sessions. In each session we used four different stimuli per probe class (linear, logarithmic, quadratic and bird vocalization sweeps) with two upward and two downward sweeps per class (or two probe stimulus pairs per probe class). The upward and downward sweep of each probe pair covered exactly the same frequency range. The pure-tone probe stimuli for each daily session were selected so that each session contained 2 'low frequency' and 2 'high frequency' linear, logarithmic and quadratic sweeps. For the first 5 sessions of the experiment, only pure-tone probe stimuli with a bandwidth of 1.6 octaves were used, whereas for the second 5 sessions stimuli with 1 and 2 octaves were used (6 of each in each session).

Each session consisted of an average of 577 completed pseudo-randomized trials for crow O and 566 completed trials for crow G. Of those, the familiar training sample stimuli were presented in 88% of the trials and probe sample stimuli were presented pseudo-randomly in the other 12% of the trails. A small proportion of probe stimuli prevented the crows to learn response patterns for those stimuli. Familiar training sample stimuli as well as probe sample stimuli were always followed by the same familiar test stimuli also used for training (see 'Training stimuli'). In either case, the crows were rewarded for every correct response to a match to promote category maintenance. Only responses to the first presentation of each unique probe stimulus were analyzed. During this first presentation of the probe stimulus, the crows were not able to learn a 'correct' response but had to infer category membership based on their previous knowledge acquired with training stimuli.

Data analysis

The percent correct responses, i.e. the number of correct trials divided by the total number of completed trials, was calculated as a measure of behavioral performance. Performance was calculated separately for up- and downward sweeping training sample stimuli and the classes of probe stimuli. To assess transfer of upward and downward FM categories, only the first trial for each unique probe FM stimulus was included. Probe trial performance therefore quantified the percentage of correctly answered first probe stimuli. This ensured that the crows could not learn how to respond to probe trials but relied on transferring their categorical perception.

Error types: The only type of error possible in the match condition is a type2-error (crow does not respond to match) because the trial ends after presentation of test1 (match). In the nonmatch condition, the crows only made type1-errors (false alarms; crow responds to nonmatch) because the crows always responded to either test1 or test2 (with the exception of a single trial across all sessions). The percent correct performance for match and nonmatch conditions separately therefore indicate all possible types of errors the crows made.