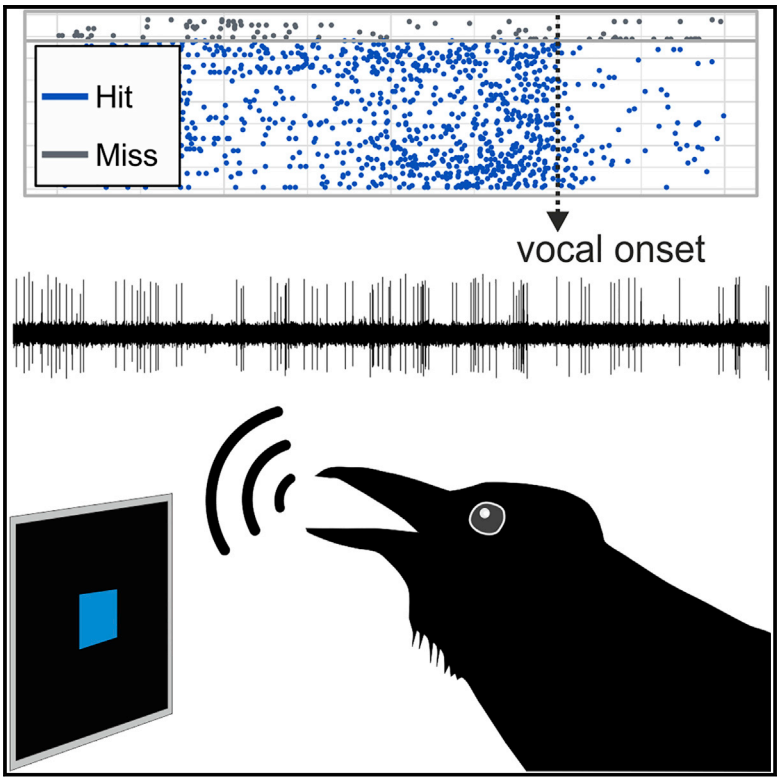
Report

Cell Reports

# Neural correlates of cognitively controlled vocalizations in a corvid songbird

## Graphical abstract




## Authors

Katharina F. Brecht,
Stephanie Westendorff, Andreas Nieder

## Correspondence

andreas.nieder@uni-tuebingen.de

## In brief

In crows trained to vocalize on command, Brecht et al. show that preparatory activity of neurons in the endbrain area nidopallium caudolaterale specifically predict the production of volitional vocalizations. This finding demonstrates that the songbird's vocal system can be controlled by executive functions.

## Highlights

- NCL neurons specifically predict whether crows produce a volitional vocalization
- Neuronal responses discriminate between volitional and task-unrelated vocalizations
- Spontaneous activity modulations influence neurons and bias vocalization decisions
- Executive control in crows can control the songbird's vocal system

Brecht et al., 2023, Cell Reports *42*, 112113
March 28, 2023 
https://doi.org/10.1016/j.celrep.2023.112113

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

# Neural correlates of cognitively controlled vocalizations in a corvid songbird

Katharina F. Brecht,[1] Stephanie Westendorff,[1] and Andreas Nieder[1,2,*]
[1]Animal Physiology Unit, Institute of Neurobiology, University of Tübingen, 72076 Tübingen, Germany
[2]Lead contact
*Correspondence: andreas.nieder@uni-tuebingen.de
https://doi.org/10.1016/j.celrep.2023.112113

## SUMMARY

The neuronal basis of the songbird's song system is well understood. However, little is known about the neuronal correlates of the executive control of songbird vocalizations. Here, we record single-unit activity from the pallial endbrain region "nidopallium caudolaterale" (NCL) of crows that vocalize to the presentation of a visual go-cue but refrain from vocalizing during trials without a go-cue. We find that the preparatory activity of single vocalization-correlated neurons, but also of the entire population of NCL neurons, before vocal onset predicts whether or not the crows will produce an instructed vocalization. Fluctuations in baseline neuronal activity prior to the go-cue influence the premotor activity of such vocalization-correlated neurons and seemingly bias the crows' decision to vocalize. Neuronal response modulation significantly differs between volitional and task-unrelated vocalizations. This suggests that the NCL can take control over the vocal motor network during the production of volitional vocalizations in a corvid songbird.

## INTRODUCTION

Vocalizations are essential for communication in most vertebrate species, including our own.[1] As vocal learners, songbirds possess a most elaborate vocal communication systems that shows many parallels with human speech.[2,3] Over the last decades, sensorimotor processes of avian vocalizations have been studied extensively. Consequently, we now have a detailed understanding of the neuronal mechanisms responsible for perceiving, learning, and producing complex vocal output in songbirds.[4–7] These complex vocal structures are learned and produced with the help of a set of distinct songbird-specific brain nuclei. This vocal system allows songbirds to communicate in a very robust and almost automatic fashion in response to affective stimuli, such as mating partners, or endogenous states influenced by hormones.[8]

A crucial characteristic of complex songbird communication is the ability to control the emission of communicative signals, that is, to use them in a cognitively controlled manner. Some recent behavioral studies showed that songbirds can control aspects of their vocalizations[9,10] or temporally adjust them to social partners.[11,12] Moreover, budgerigars can be trained to make particular vocalizations in an operant setup[13,14] We have previously demonstrated that carrion crows, songbirds of the corvid family that possess the standard set of song nuclei,[15] can be trained to vocalize in response to non-hedonic cues presented on a screen and to refrain from vocalizing to other cues.[16] This result suggests that crows can use vocalizations in a volitional manner, beyond affective responses to arousing stimuli.[17] While the neuronal mechanisms of the sensorimotor aspects of songbird vocalizations are well understood, how songbirds may gain control over their vocal output remained elusive.

To assess volitional control of crow vocalizations, we adopted an operational definition from clinical neurology where the distinction between volitional and affective vocalizations has long been recognized for the diagnosis of the type of orofacial paralysis:[18–22] First, volitional vocalizations need to be uttered in response to an arbitrary instruction stimulus that is neutral in its value or emotional valence. Second, vocalizations need to be uttered in a manner that is temporally contingent to the instruction stimulus. Third, vocalizations need to be produced reliably after the presentation of the instructive stimulus and withheld in its absence.[16,23]

In the present study, we explored how the neuronal correlates of cognitive vocal control in crows demonstrate sophisticated cognitive behaviors.[24,25] Cognitive control in birds is associated with the nidopallium caudolaterale (NCL) located in the avian telencephalon,[26,27] which is often called the "avian prefrontal cortex."[28,29] As the highest associative integration center,[30] the NCL receives and integrates information across sensory modalities,[31,32] groups sensory input according to categories and rules,[33–35] evaluates information according to reward value,[36] maintains information online in working memory,[37,38] and controls motor structures by representing the initiation of planned movements.[39] To investigate the role of the NCL in eliciting volitional vocalizations, we recorded single-neuron activity from the NCL of crows exerting cognitive control over their vocalizations.

## RESULTS

Two crows performed a computerized visual detection task where they had to respond to the presentation of a specific cue ("go-cue") with a vocalization and to refrain from vocalizing

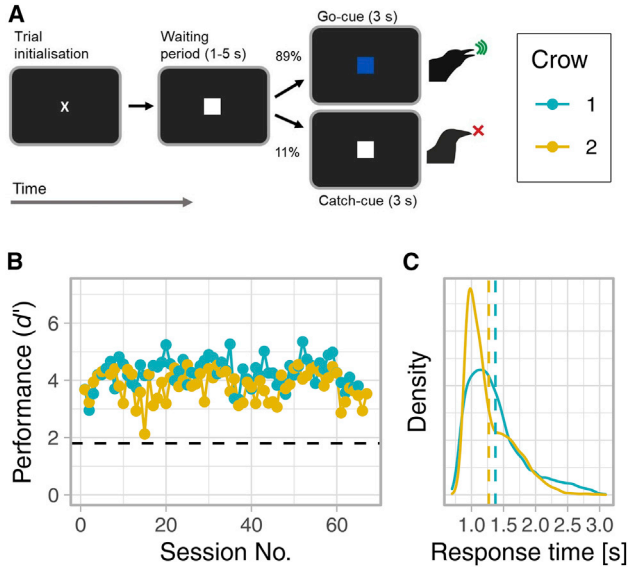



**Figure 1. Behavioral protocol and performance**

(A) Crows were trained on a self-initiated visual detection task with vocalizations as response. In 89% of the cases, the color of the square changed to blue or red (50% of trials each, go trials), indicating the go period in which the crows could vocalize to obtain a reward. In the other 11% of the cases (catch trials), the white square remained, and the crow was required to refrain from vocalizing.

(B) Performance measured as *d*' (log-linear adjusted) for both crows across all recording sessions. Dashed line denotes *d*' = 1.8, which corresponds to successful signal detection.

(C) Response time (vocal onset) distribution for successful vocalizations for both crows. Dashed lines indicate the mean response time.

when another cue ("catch cue") was presented (Figure 1A). The crows' responses in the two different trial conditions were categorized according to signal detection theory.[40] A correct vocalization within 3 s after go-cue was defined as "hit." If the crows did not vocalize in response to the go-cue, the trial was counted as "miss." Correspondingly, a vocalization in a catch-trial was defined as a "false alarm (FA)," whereas withholding vocalization in a catch-trial was counted as "correct rejection (CR)." We had reported in a previous study that the crows have volitional control over their vocalizations in this behavioral protocol.[16] To investigate whether and how the NCL controls volitional vocalizations, we now concurrently recorded single neuron activity from the behaving crows' NCL.

Crow 1 participated in 65 daily sessions, and crow 2 in 63 sessions. Crows vocalized on average 255 ± 61 and 152.99 ± 53 times per session, respectively, and reliably did so in response to the presentation of a go-cue in go trials (Figure 1B). The crows never vocalized during catch trials, indicating that they were not simply vocalizing after a certain time had elapsed but understood the task contingencies. To evaluate the crows' performance, we calculated the sensitivity measure *d*'. In all sessions, the *d*' was significantly above the threshold value of 1.8 (Fisher-Pitman exact permutation test, $p < 0.001$) for both crows. Crow 1 (m = 1.371 s, SD = 0.445 s) showed longer vocal reaction times than crow 2 (m = 1.268s, SD = 0.355 s, Kolmogorov-Smirnov test, D = 0.143, $p < 0.01$) (Figure 1C).

## Neuronal responses prior to cued vocalization

To investigate whether and how the crows' brains control volitional vocalizations, we recorded from a total of 287 single neurons (166 in crow 1 and 121 in crow 2) from the behaving crows' NCL (for inclusion criteria, see STAR Methods). To identify neurons that showed significant premotor response modulation, we compared the average firing rates (after cue onset but prior to vocal output) during hit trials with those elicited during miss trials. Overall, 16% (47/287) of all neurons showed a significant difference in firing rates (Wilcoxon rank-sum test; $p < 0.05$). These neurons were termed "vocalization-correlated neurons." Vocalization-correlated neurons were found in both crows, with crow 1 exhibiting more vocalization-correlated cells (39/166, 23.5%) than crow 2 (8/121, 6.6%, chi-squared test, $\chi^2 = 9.732$, $p = 0.002$, Figure 2). The example neurons in Figures 2A and 2C showed significantly higher firing rates prior to the onset of volitional vocalizations (hit trials) compared with miss trials for crow 1 and 2, respectively. In contrast, the neurons displayed in Figures 2B and 2D exhibited lower discharges in hits relative to misses.

These response patterns were observed across the population of vocalization-correlated neurons. Overall, 38% of vocalization-correlated neurons increased activity during hit trials compared with miss trials (selectively increasing neurons, n = 18), whereas the remaining 62% showed significantly reduced activity during hit trials (selectively decreasing neurons, n = 29). The average normalized mean firing rates of selectively increasing and selectively decreasing vocalization-correlated neurons are shown in Figures 2E and 2F. The firing rate differences (Figures 2E and 2F; hit versus miss) visible very early in the trial indicate systematic differences in firing rates in the baseline activity that we investigated next.

We explored whether baseline activity fluctuations prior to the onset of the go-cue would predict the likelihood that the crows produced a hit or rather missed a vocalization. If so, vocalization-correlated neurons showing an increased firing rate prior to the vocalizations in hit compared with miss trials should, on average, exhibit higher baseline firing rates. Conversely, vocalization-correlated neurons with decreased firing rate prior to the vocalizations in hit compared with miss trials should, on average, exhibit lower baseline firing rates. Indeed, we found that the interaction of trial type (hit versus miss) and type of neuron (increasing versus decreasing preparatory activity) had a significant effect on firing rate prior to go-cue onset ($F(1.45) = 128.83$, $p < 0.01$, aligned rank transform ANOVA). For neurons with increased preparatory firing rates, the average baseline firing rate prior to go-cue onset in hit trials (average firing rate = 5.26 Hz) was significantly higher than in miss trials (average firing rate = 1.94 Hz, W = 171, $p < 0.001$, Wilcoxon signed rank test; Figure 2G). For neurons with decreased preparatory firing rates, the average baseline firing rate prior to go-cue onset in hit trials (average firing rate = 1.06 Hz) was significantly lower than in miss trials (average firing rate = 3.4 Hz, W = 435, $p < 0.001$, Wilcoxon signed rank test; Figure 2H). Thus, baseline activity fluctuations correlated with the crows' vocalization-correlated neurons as well as the crows' vocal behavior.

## Dynamics of neuronal activity

To explore the temporal evolution of activation of the neurons that had been tested as significantly vocalization correlated,

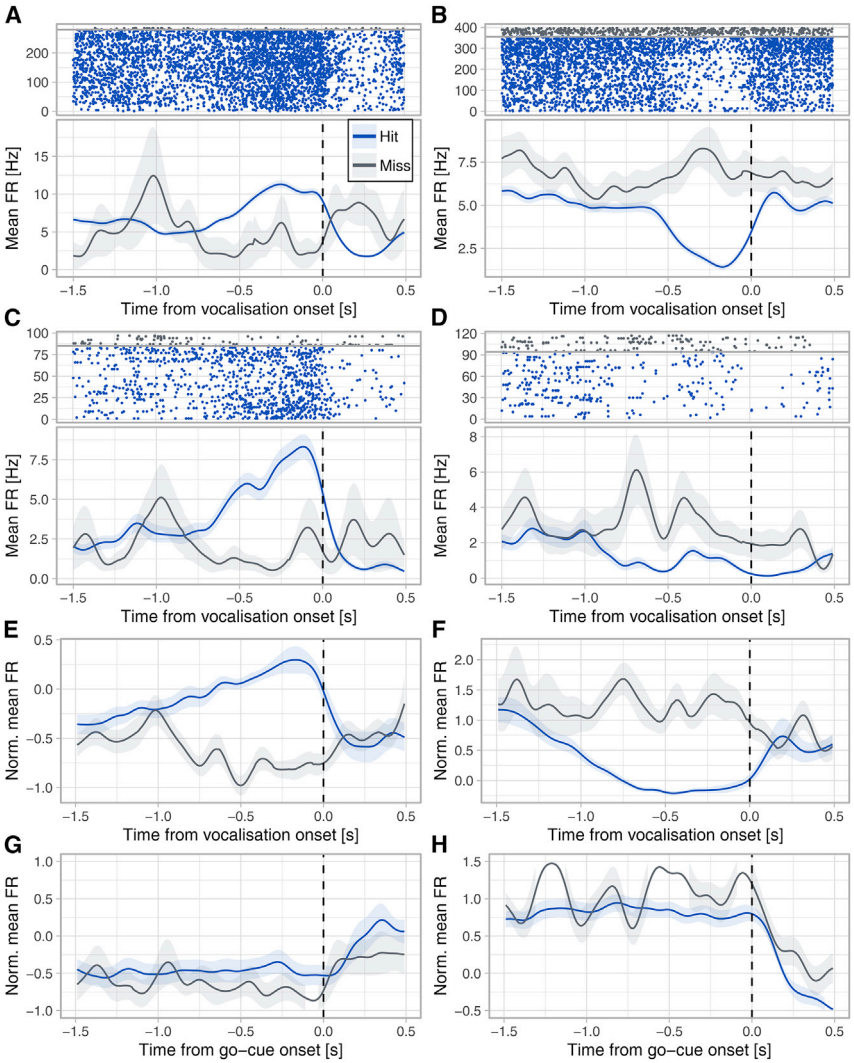



**Figure 2. Neuronal response prior to volitional vocalization**

(A) Response of example neuron from crow 1 showing a significant increase of neuronal activity during trials with vocalizations in response to the go-cue, compared with trials without vocalization (miss trials). Upper panel shows dot raster plots, with blue dots representing single action potentials during vocalization trials, and gray dots representing action potentials during miss trials. Each line represents one trial. Lower panels show the corresponding peri-stimulus histogram, averaged over trials. Shaded areas indicate standard error.

(B) Example neuron from crow 1 with a decrease in activity. Layout as in (A).

(C and D) Example neurons from crow 2 with increasing and decreasing activity, respectively.

(E and F) Averaged and normalized activity of vocalization-correlated neurons for all neurons with increasing (E) and decreasing (F) activity prior to vocal onset in hit trials.

(G and H) Averaged and normalized baseline activity of vocalization-correlated neurons with increasing (G) and decreasing (H) activity prior to go-cue onset correlated with the impending vocalization.

we determined the time points of selective firing rate differences between hit-miss differences using a sliding-window analysis. To that aim, we computed effect sizes for each neuron from a sliding Wilcoxon rank-sum test based on firing rate distributions. We used receiver operating characteristic curve (AUC) to estimate effect sizes for time windows in which the Wilcoxon rank-sum test was significant. Values closer to 1 or 0 indicate better separation of the two distributions, and 0.5 indicates complete overlap. In this case, values closer to 1 translate to higher firing rates in hit trials, whereas values closer to 0 to higher firing rates in miss trials (Figure 3A).

We further explored whether and how the entire population of recorded NCL neurons, irrespective of any response preferences, encoded hit and miss trials. Therefore, we performed a principle component analysis on a population of pseudo-simultaneously recorded neurons (n = 58) for which a sufficient number of hit and miss trials could be recorded. This approach extracts trajectories from the spiking activity of a neuronal population in individual trials. Such trajectories reflect the instantaneous firing rates of the respective neuronal population as they evolve over time. Figure 3B depicts average population trajectories for hit and miss trials in a space defined by the top three most meaningful dimensions. To evaluate the temporal evolution of population activity and to different trial types, we measured Euclidian distances between trial trajectories corresponding to the hit and miss trials. The analysis shows that the distance between hit and miss trials prominently increased in the 500-ms interval before vocal onset (Figure 3C). Based on Euclidian distances, a classifier was able to significantly predict hit from miss trials in a time window starting 560 ms before vocal onset (randomization test, see STAR Methods), (Figure 3D). This result shows that information about the preparation of volitional vocalizations is not only present in individual selective neurons but also across the entire population of NCL neurons.

## Volitional compared with task-unrelated vocalizations

To explore whether vocalization-correlated activity was specific for volitional vocalizations rather than a signature of any vocalization the crows may have elicited, we contrasted vocalization-correlated activity during hits with the task-unrelated vocalizations the crows made unrelated to the task contingencies. Because the crows vocalized almost exclusively to the go-cue during the task, overall, there were very few instances of task-unrelated vocalizations (similar to what we reported in Brecht et al.[16]).

In 29 out of the 47 recorded vocalization-correlated neurons, the crows produced a sufficient number ($\geq$3) of task-unrelated vocalizations. For these 29 neurons, we compared the mean

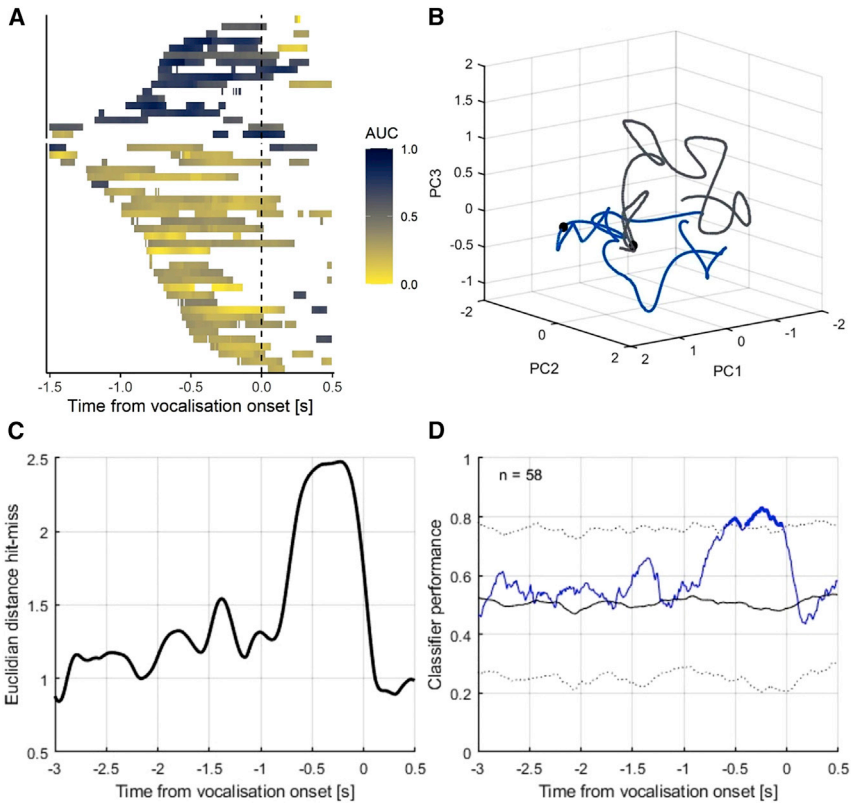



**Figure 3. Dynamics of neuronal activity**

(A) Time course of the activity of vocalization-correlated cells prior to vocalization onset. Each row represents one cell. Cells are ordered based on activity onset and divided into selectively increasing (upper panel) and selectively decreasing (lower panel). Effect sizes (calculated as AUC) of a sliding Wilcoxon rank-sum test comparing hit and miss trials are binned in quantiles for illustrative purposes. Vocalization onset is represented by the black dashed line. Time 0 ms represents vocal onset.

(B) Activity trajectory of the entire population of recorded NCL neurons (irrespective of response preferences) in hit (blue) and miss (gray) trials in the space spanned by the first three principle components. Black dots represent vocalization onset.

(C) Euclidian distance between the trajectories of hit and miss trials shown in (B).

(D) The performance of a classifier trained on the trajectory data of the entire population of recorded NCL neurons significantly predicted impending hit and miss trials (blue). The black solid lines represent the mean of a distribution with shuffled trial labels. Dashed lines depict the significance threshold, i.e., the 2.5th and 97.5th percentile of the shuffled distribution.

firing rate for volitional and task-unrelated vocalizations. Across the neuron population, the average firing rate 1,000 ms before a task-unrelated vocalization ($m_{\text{average firing rate}}$ = 5.15) was significantly different from that prior to a volitional vocalization ($m_{\text{average firing rate}}$ = 2.57, $Z = -2.78$, p = 0.0055, Wilcoxon rank-sum test). In addition, we compared the discharges to volitional and task-unrelated vocalizations for each individual cell. In 15 of these 29 neurons, the comparison of firing rates 1,000 ms prior to vocalization was significantly different (Wilcoxon rank-sum test; see example neuron shown in Figure 4A); that is, 52% of the vocalization-correlated neurons for which we were able to record task-unrelated vocalizations showed significant differences between volitional and task-unrelated vocalization in firing rate prior to the vocalization onset on the single-cell level. Figure 4B shows, equivalent to Figure 3A, the effect sizes as AUC values in a sliding-window analysis for all cells that showed a significant difference in firing rates between task-unrelated and volitional vocalizations. Figure 4C depicts the normalized activity prior to response onset for selectively decreasing cells (volitional vocalization < task-unrelated vocalization, $p < 0.05$, n = 13). Selectively increasing cells (volitional vocalization > task-unrelated vocalization, $p < 0.05$, n = 2) are not shown.

Since volitional and task-unrelated vocalizations on average showed subtle differences in acoustic features, we explored whether the observed firing rate differences between volitional and task-unrelated vocalizations might reflect vocalization acoustics rather than task context. To that aim, we repeated the comparison of individual neurons' (n = 29) firing rates for volitional versus task-unrelated vocalization only for trials that were matched in acoustic features. In other words, we compared individual neurons' activity only in trials in which the acoustics of the volitional and task-unrelated vocalizations were most similar. Figure 4E shows that the vocalizations with matched acoustic parameters were very similar and thus strongly overlapping in a three-dimensional acoustic parameter space; the Euclidian distances between the matched vocalizations were significantly smaller than those of randomly paired vocalizations ($p < 10^{-13}$, randomization test).

If differences in vocalization acoustics would drive the neuronal activity differences, the firing rate differences between volitional and task-unrelated vocalizations should vanish for acoustically matched trials. However, this is not what we observed. We first matched the duration, maximum frequency, and Wiener entropy of volitional and task-unrelated vocalizations (Figure S1). After matching, no evidence for a difference in Wiener entropy (median cued_voc = –1.23 ± 0.02, task-unrelated_voc = −1.21 ± 0.02, p = 0.15, Wilcoxon rank-sum test, n = 943; Cohen's d: 0.06) (Figure S1B) or the maximum frequency (median cued_voc = 574 ± 6 Hz, task-unrelated_voc = 579 ± 7 Hz, p = 0.43, Cohen's d: 0.04) (Figure S1C) was observed. A difference with a small effect size (Cohen's d: 0.13) occurred only between the duration in cued vocalizations (median = 291 ± 3 ms) and task-unrelated vocalizations (300 ± 4 ms; p = 0.002, Wilcoxon rank-sum test) (Figure S1A). The median of normalized differences between task-unrelated and matched vocalizations (expressed in multiples of the pooled standard deviation) across all three parameters was 0.52, which was only half the magnitude compared with the values of between 0.87 and 1.02 for randomly

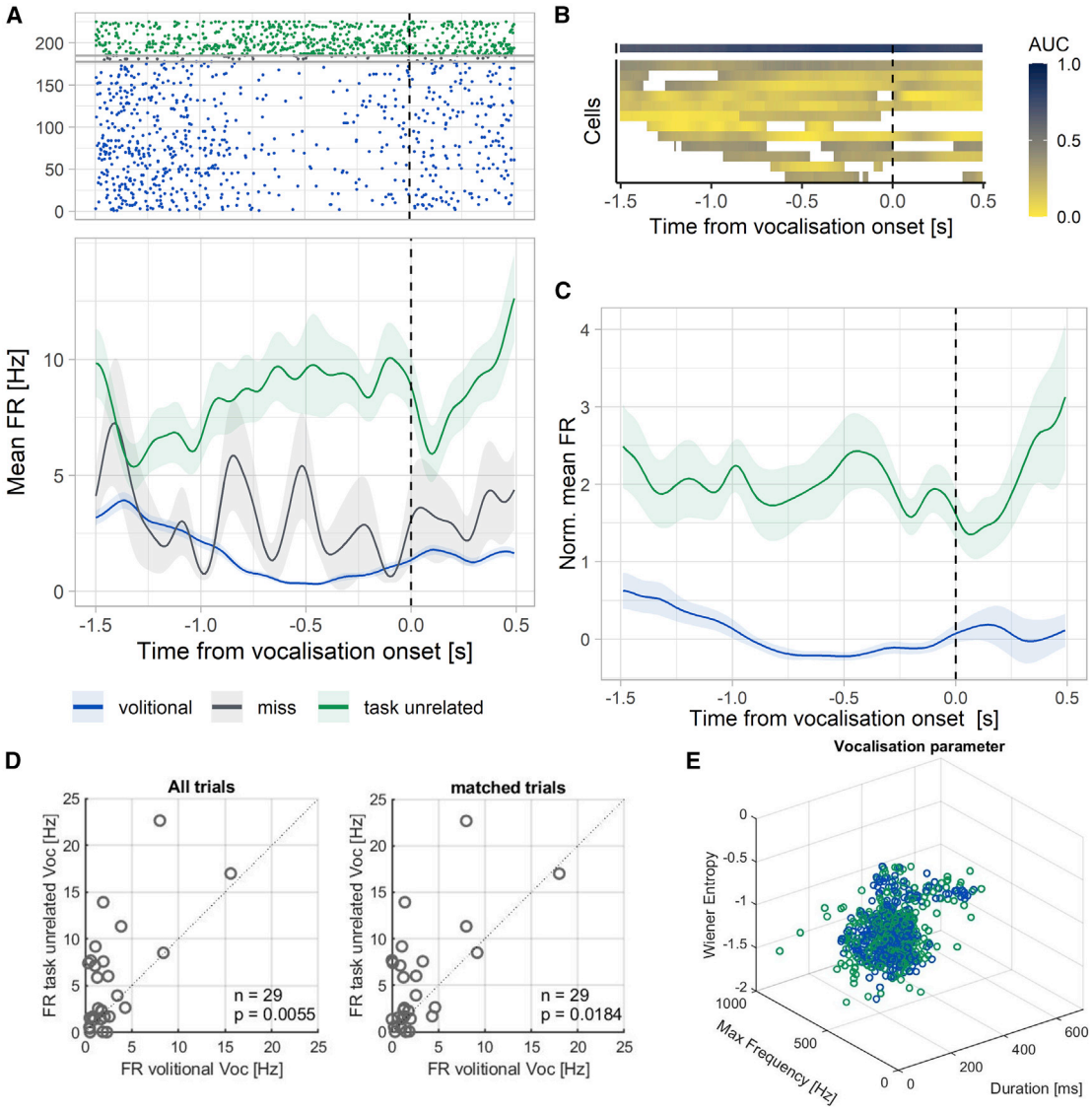



**Figure 4. Neuronal responses prior to volitional compared with task-unrelated vocalizations**

(A) Response of an example vocalization-specific neuron showing a significant decrease of neuronal activity prior to volitional vocalizations compared with task-unrelated vocalizations. Top: dot-raster histogram; bottom: spike-density histogram. Layout as in Figure 2A.

(B) Effect sizes (calculated as AUC) calculated for vocalization-correlated neurons. Each row represents one cell. Cells are ordered based to activity onset and divided into selectively increasing (upper panel) and selectively decreasing (lower panel). Vocalization onset is represented by the black dashed line.

(C) Averaged and normalized activity of vocalization-correlated neurons that decreased in activity prior to volitional vocalizations.

(D) Left: firing rate in trials with volitional versus task-unrelated vocalizations if all trials for the volitional vocalizations are considered. Each dot represents the firing rate in one cell in the 1 s before the onset of the vocalization.

Right: firing rate in trials with volitional versus task-unrelated vocalizations if only trials are considered in which the vocalization parameter of the volitional vocalizations were matched to the parameters of the task-unrelated vocalizations.

(E) Distribution of volitional (blue) and task-unrelated (green) vocalizations in the three-dimensional parameter space of analyzed vocalization parameters (only matched trials). Each dot represents one vocalization.

chosen pairs of vocalizations (Figure S1D). Overall, the matched cued and task-unrelated vocalizations conformed well.

Importantly, the firing rate differences persisted even when the acoustics of the volitional and task-unrelated vocalizations were matched ($Z = -2.36$, $p = 0.018$, Wilcoxon rank-sum test) (Figure 4D). The magnitudes of the firing rate differences between volitional and task-unrelated vocalizations for matched vocalizations (2.79 Hz versus 5.15 Hz) were almost identical to those observed when the vocalizations were acoustically unmatched (2.57 Hz versus 5.15 Hz). This finding suggests that the observed preparatory activity differences between volitional and task-unrelated vocalizations were independent from the acoustic parameters of the impending vocalizations, and thus reflect volitional versus task-unrelated vocal preparation.

## DISCUSSION

Our electrophysiological results provide a first answer to the question of how songbirds may gain cognitive control over their vocal output. We report a neuronal correlate of crows' volitional vocalization initiation in the NCL, a high-level associative endbrain structure in birds that operates at the apex of the avian pallial hierarchy.[41] The activity of single vocalization-correlated NCL neurons prior to vocal onset predicted whether or not the crows would produce an instructed vocalization. The differences we observed in the proportions of vocalization-correlated neurons between the two crows may be due to slight differences in NCL recording sites. Importantly, firing rates were significantly different between volitional and task-unrelated vocalizations. This is evidence that this neuronal activity was not just signaling the preparation of any vocal output, but the voluntary initiation of vocalizations specifically.

### NCL connections to the vocal motor network

Crows possess all songbird-typical song-related structures.[15] Currently, however, it is unknown how and at which processing level NCL may take control over the vocal motor network. Considering that the NCL is a telencephalic association area, a direct monosynaptic control of phonatory motoneuron pools seems highly unlikely. We speculate that NCL output projects to other telencephalic territories involved in vocal output and motor control.

One candidate recipient of NCL projections is the premotor song nucleus HVC that operates at the apex of the song motor system.[42] However, whether the HVC receives projections from higher association brain areas is currently unknown. Interestingly, both HVC and NCL originate from the same larger nidopallial territory. The HVC has even been suggested to constitute a songbird specialization of the NCL.[43] Besides this anatomical similarity, there are also functional commonalties between HVC and NCL as both structures control and initiate learned sequences[42,44]

In addition, NCL projects to parts of the arcopallium[45,46] and may thereby mediate cognitive control over vocal output. In songbirds, the arcopallium is involved in general motor generation,[47,48] but more specifically also in song generation via the vocal motor nucleus (RA). Moreover, the NCL has been identified as part of a recurrent and inter-hemispheric circuit within a cortico-basal ganglia pathway necessary for vocal learning.[46] Thus, in addition to the NCL's general role in multimodal learning-related behaviors,[32,49] it might also mediate cognitive control over vocal behavior via connections to dedicated song nuclei. Such a connection between the associative NCL and song-related structures could also explain how multimodal social signals impact proper vocal learning and perhaps communication signals in general in songbirds.[50] Ultimately, the connectivity patterns and functional links between the NCL and other vocal-related brain nuclei need to be investigated, both with anatomical explorations and physiological manipulations such as electrical stimulation.[51,52]

The proposed interplay of the NCL with vocal-related brain structures may constitute an evolutionary novelty of songbirds. This is because the caudal telencephalon, specifically the caudal nidopallium and arcopallium, is very similar in organization, architecture, and dopaminergic innervation,[53–56] but it has evolved strikingly differently compared with pigeons and chickens.[57] Indeed, it has been suggested that vocal learning brain pathways evolved out of general motor-learning pathways that control complex, learned non-vocal behaviors.[5,48,58–60] Such ancestral motor control circuits presumably also had NCL input to them to control other motor behaviors. This reorganization of the caudal telencephalon seems to be unique for members of the oscines (songbirds) and may have emerged with the rise from their last common ancestor at least 40 million years ago.[61]

### NCL function in vocal control parallels primate PFC

The current findings also point to a further interesting functional parallel between the NCL in crows and the prefrontal cortex (PFC) in primates. These two high-level associative telencephalic brain areas evolved independently through convergent evolution. As a result, the neuroanatomy of both brain areas is strikingly different. For instance, the NCL shows no layered organization that is characteristic for the neocortex. Despite such major differences, the NCL and the PFC seem to adopt similar cognitive functionality.[26,28]

Using the exact same behavioral protocol, macaque monkeys have been shown to also be able to cognitively control their vocalizations.[62–64] The two tested monkeys showed even longer reaction times (RTs; on average 1.53 s and 1.64 s, respectively) in the vocal task compared with the relatively quick crows (1.27 and 1.37 s, respectively). Compared with volitional hand or eye movements, vocal RTs are considerably slower. This is likely a consequence of the complex preparatory coordination of respiratory, orofacial, and vocal (laryngeal/syringeal) muscles required for instructed vocalizations.[65] In agreement with such long RTs, electrical stimulation of vocal-related cortical areas in monkeys precedes vocal onset by as long as 1,000 to 2,900 ms.[66,67] These stimulation latencies are in stark contrast to those reported for eye movements, which can be elicited 20 to 60 ms after frontal eye field stimulation.[68] During this long-lasting preparatory period, we detected relatively early and long-lasting (several hundred milliseconds) preparatory activity of vocalization-correlated neurons in crows. Similar long-latency preparatory activity of up to 1 s has been reported for cortical neurons of vocalizing monkeys.[63,69–72]

During recordings in the macaque dorsolateral PFC, 15% of the randomly recorded neurons were classified as vocalization-correlated neurons.[63] This proportion of vocalization-correlated neurons in primate PFC is almost identical to the 16% of vocalization-correlated neurons we found in the crow NCL. Also similar to the crow NCL, about half of the vocalization-correlated PFC neurons showed significant differences in pre-vocal discharge rates between volitional and task-unrelated vocalizations.[63] In both monkeys and crows, this finding suggests a strong involvement of the PFC and NCL, respectively, in the initiation of volitional vocalizations.[72]

### Baseline fluctuations

We found that fluctuations in baseline neuronal activity immediately preceding go-cue presentation influenced the premotor activity of vocalization-correlated neurons and biased the crows' decision to vocalize (hit) or to withhold a vocalization (miss).

The baseline fluctuations were specifically predictive of the behavioral consequences for the two classes of (increasing or decreasing) vocalization-correlated neurons: low baseline activity in neurons showing increasing premotor activity - and high baseline activity in neurons exhibiting decreasing premotor activity - correlated with more misses. If baseline activity of the class of neurons showing increasing premotor activity was low prior to the go-cue, the crows were prone to misses. Conversely, if baseline activity of the class of neurons showing decreasing premotor activity was high prior to the go-cue, the crows were also prone to misses.

These correlations of baseline activity in the crow mirror findings in decision-making mammals. Variations in activity of cortical neurons just before stimulus onset can be predictive of the animals' subsequent perceived stimulus.[73–77] Although fluctuations in pre-cue activity were related to vocal production performance in our crows, the origin of the fluctuations is unknown. It is likely that the level of baseline activity is related to events occurring on previous trials.[78] Several studies in the mammalian cortex reported so-called sequential biases in which current choices can depend on previous history of stimuli, rewards, and choices even in the conditions where such dependence is not task relevant.[79–82] Similar sequential biases could be at work in vocalizing crows.

### Balanced excitation and inhibition during volitional motor control

We found many vocalization-correlated neurons in the crow NCL that reduced their activity in hit trials. In a previous study in vocalizing monkeys, we similarly reported that about half of the vocalization-correlated neurons in motor-related frontal lobe areas decreased their activity; this was observed not only for the preparation of volitional vocalizations but even for hand movements.[73] At first glance, this finding may sound incompatible with preparatory motor activity requiring activation of vocal motor circuits. However, motor circuit operations not only rely on excitation but also on balancing inhibition. For instance, inhibitory interneurons in mouse motor cortex increased their firing in response to a movement cue as well as the onset of reaching, suggesting that inhibitory interneurons participate in voluntary movement execution by inhibiting excitatory projection neurons.[83] Moreover, activation of different types of inhibitory interneurons can reduce or enhance locomotion.[84] Overall, these data point to distinct roles of inhibitory neurons and suppressed activity in motor circuits involved in voluntary skilled movements; these findings indicate that balanced excitation and inhibition is crucial for proper motor output.[85]

### Limitations of the study

To minimize the influence of purely acoustic vocal differences on the observed neuronal differences between volitional and task-unrelated vocalisations, we matched important acoustic parameters. However, we cannot be certain that these three equalized parameters represent the most salient features to the crows. Aspects of the observed neuronal differences between volitional and task-unrelated vocalizations may therefore still relate to the acoustics of these types of vocalizations made by the crows.

Given that the NCL is also involved in the planning and execution of volitional head-beak actions, the specificity of activity in NCL for volitional vocalisations needs further investigation. Monkey data at least suggest separate populations of neurons for volitional motor preparation. When monkeys were trained to execute two types of volitional actions during the same recordings, vocalizations or hand movements, neurons indeed differentiated between the volitional initiation of vocal and manual acts.[73] Given that the NCL is considered the avian equivalent of the mammalian PFC, the premotor activity we describe in the current paper in crows is very likely specific for vocal output rather than a general reflection of volitional acts. Of course, despite the data showing a strong involvement of the NCL in volitional vocalizations, our findings cannot rule out a putative additional role in task-unrelated and spontaneous vocalizations.

The current findings are based on correlations of neuronal activity with the crows' behavior. To move from correlation to causation, measuring the vocal consequences of manipulating NCL would be a great asset. More insights into the connectivity patterns of the corvid NCL, in particular those to the vocal production system, are required to predict the consequences of neuronal manipulations. Since NCL perturbations in birds result in several complex deficits, more differentiating behavioral protocols are needed to break down the expected complex deficits into specific cognitive components.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Subjects
- METHOD DETAILS
  - Set-up
  - Behavioral protocol
  - Surgery and neurophysiological recordings
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral analysis
  - Neuronal analysis

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112113.

### ACKNOWLEDGMENTS

This research was supported by DFG grant BR 308/1-1.

### AUTHOR CONTRIBUTIONS

K.F.B. and A.N. designed the experiment, K.F.B. recorded the data, K.F.B., S.W., and A.N. analyzed the data, K.F.B., S.W., and A.N. wrote the paper, and A.N. supervised the study.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 19, 2022
Revised: January 13, 2023
Accepted: January 28, 2023

## REFERENCES

1. Rohrmeier, M., Zuidema, W., Wiggins, G.A., and Scharff, C. (2015). Principles of structure building in music, language and animal song. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20140097.
2. Doupe, A.J., and Kuhl, P.K. (1999). Birdsong and human speech: common themes and mechanisms. Annu. Rev. Neurosci. *22*, 567–631.
3. Bolhuis, J.J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. Nat. Rev. Neurosci. *11*, 747–759.
4. Farries, M.A. (2004). The avian song system in comparative perspective. Ann. N. Y. Acad. Sci. *1016*, 61–76.
5. Jarvis, E.D. (2019). Evolution of vocal learning and spoken language. Science *366*, 50–54.
6. Schmidt, M.F., and Martin Wild, J. (2014). The respiratory-vocal system of songbirds: anatomy, physiology, and neural control. Prog. Brain Res. *212*, 297–335.
7. Simonyan, K., Horwitz, B., and Jarvis, E.D. (2012). Dopamine regulation of human speech and bird song: a critical review. Brain Lang. *122*, 142–150.
8. Gahr, M. (2014). How hormone-sensitive are bird songs and what are the underlying mechanisms? Acta Acustica united Acustica *100*, 705–718.
9. Tumer, E.C., and Brainard, M.S. (2007). Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature *450*, 1240–1244.
10. Veit, L., Tian, L.Y., Monroy Hernandez, C.J., and Brainard, M.S. (2021). Songbirds can learn flexible contextual control over syllable sequencing. Elife *10*, e61610.
11. Ma, S., Ter Maat, A., and Gahr, M. (2020). Neurotelemetry reveals putative predictive activity in HVC during call-based vocal communications in zebra finches. J. Neurosci. *40*, 6219–6227.
12. Benichov, J.I., and Vallentin, D. (2020). Inhibition within a premotor circuit controls the timing of vocal turn-taking in zebra finches. Nat. Commun. *11*, 221.
13. Seki, Y., and Dooling, R.J. (2016). Effect of auditory stimuli on conditioned vocal behavior of budgerigars. Behav. Processes *122*, 87–89.
14. Osmanski, M.S., Seki, Y., and Dooling, R.J. (2021). Constraints on vocal production learning in budgerigars (Melopsittacus undulates). Learn. Behav. *49*, 150–158.
15. Kersten, Y., Friedrich-Müller, B., and Nieder, A. (2021). A histological study of the song system of the carrion crow (*Corvus corone*). J. Comp. Neurol. *529*, 2576–2595.
16. Brecht, K.F., Hage, S.R., Gavrilov, N., and Nieder, A. (2019). Volitional control of vocalizations in corvid songbirds. PLoS Biol. *17*, e3000375.
17. Liao, D.A., Zhang, Y.S., Cai, L.X., and Ghazanfar, A.A. (2018). Internal states and extrinsic factors both determine monkey vocal production. Proc. Natl. Acad. Sci. USA. *115*, 3978–3983.
18. Rinn, W.E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. Psychol. Bull. *95*, 52–77.
19. Hopf, H.C., Müller-Forell, W., and Hopf, N.J. (1992). Localization of emotional and volitional facial paresis. Neurology *42*, 1918–1923.
20. Van Lancker, D., and Cummings, J.L. (1999). Expletives: neurolinguistic and neurobehavioral perspectives on swearing. Brain Res. Brain Res. Rev. *31*, 83–104.
21. Cattaneo, L., and Pavesi, G. (2014). The facial motor system. Neurosci. Biobehav. Rev. *38*, 135–159.
22. Scott, S.K. (2022). The neural control of volitional vocal production-from speech to identity, from social meaning to song. Philos. Trans. R. Soc. Lond. B Biol. Sci. *377*, 20200395.
23. 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.
24. Hedges, S.B. (2002). The origin and evolution of model organisms. Nat. Rev. Genet. *3*, 838–849.
25. Karten, H.J. (2015). Vertebrate brains and evolutionary connectomics: on the origins of the mammalian 'neocortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20150060.
26. Güntürkün, O. (2005). The avian "prefrontal cortex" and cognition. Curr. Opin. Neurobiol. *15*, 686–693.
27. Nieder, A. (2017). Inside the corvid brain—probing the physiology of cognition in crows. Current Opinion in Behavioral Sciences *16*, 8–14.
28. Divac, I., Mogensen, J., and Björklund, A. (1985). The prefrontal "cortex" in the pigeon. Biochemical evidence. Brain Res. *332*, 365–368.
29. Mogensen, J., and Divac, I. (1982). The prefrontal "cortex" in the pigeon. Behavioral evidence. Brain Behav. Evol. *21*, 60–66.
30. Nieder, A., Wagener, L., and Rinnert, P. (2020). A neural correlate of sensory consciousness in a corvid bird. Science *369*, 1626–1629.
31. Kröner, S., and Güntürkün, O. (1999). Afferent and efferent connections of the caudolateral neostriatum in the pigeon (Columba uvia): a retro- and anterograde pathway tracing study. J. Comp. Neurol. *407*, 228–260.
32. Moll, F.W., and Nieder, A. (2015). Cross-modal associative mnemonic signals in crow endbrain neurons. Curr. Biol. *25*, 2196–2201.
33. Ditz, H.M., Fechner, J., and Nieder, A. (2022). Cell-type specific pallial circuits shape categorical tuning responses in the crow telencephalon. Commun. Biol. *5*, 269.
34. 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.
35. Veit, L., and Nieder, A. (2013). Abstract rule neurons in the endbrain support intelligent behaviour in corvid songbirds. Nat. Commun. *4*, 2878.
36. Dykes, M., Klarer, A., Porter, B., Rose, J., and Colombo, M. (2018). Neurons in the pigeon nidopallium caudolaterale display value-related activity. Sci. Rep. *8*, 5377.
37. Veit, L., Hartmann, K., and Nieder, A. (2014). Neuronal correlates of visual working memory in the corvid endbrain. J. Neurosci. *34*, 7778–7786.
38. Rinnert, P., Kirschhock, M.E., and Nieder, A. (2019). Neuronal correlates of spatial working memory in the endbrain of crows. Curr. Biol. *29*, 2616–2624.e4.
39. Rinnert, P., and Nieder, A. (2021). Neural code of motor planning and execution during goal-directed movements in crows. J. Neurosci. *41*, 4060–4072.
40. Green, D.M., and Swets, J.A. (1966). In Signal Detection Theory and Psychophysics, Repr., ed. (Peninsula Publ).
41. Leutgeb, S., Husband, S., Riters, L.V., Shimizu, T., and Bingman, V.P. (1996). Telencephalic afferents to the caudolateral neostriatum of the pigeon. Brain Res. *730*, 173–181.
42. Yu, A.C., and Margoliash, D. (1996). Temporal hierarchical control of singing in birds. Science *273*, 1871–1875.
43. Farries, M.A. (2001). The oscine song system considered in the context of the avian brain: lessons learned from comparative neurobiology. Brain Behav. Evol. *58*, 80–100.

44. Helduser, S., Cheng, S., and Güntürkün, O. (2013). Identification of two forebrain structures that mediate execution of memorized sequences in the pigeon. J. Neurophysiol. *109*, 958–968.

45. Mandelblat-Cerf, Y., Las, L., Denisenko, N., and Fee, M.S. (2014). A role for descending auditory cortical projections in songbird vocal learning. Elife *3*, e02152.

46. Paterson, A.K., and Bottjer, S.W. (2017). Cortical inter-hemispheric circuits for multimodal vocal learning in songbirds. J. Comp. Neurol. *525*, 3312–3340.

47. Bottjer, S.W., Brady, J.D., and Cribbs, B. (2000). Connections of a motor cortical region in zebra finches: relation to pathways for vocal learning. J. Comp. Neurol. *420*, 244–260.

48. Feenders, G., Liedvogel, M., Rivas, M., Zapka, M., Horita, H., Hara, E., Wada, K., Mouritsen, H., and Jarvis, E.D. (2008). Molecular mapping of movement-associated areas in the avian brain: a motor theory for vocal learning origin. PLoS One *3*, e1768.

49. Braun, K., Bock, J., Metzger, M., Jiang, S., and Schnabel, R. (1999). The dorsocaudal neostriatum of the domestic chick: a structure serving higher associative functions. Behav. Brain Res. *98*, 211–218.

50. Carouso-Peck, S., and Goldstein, M.H. (2019). Female social feedback reveals non-imitative mechanisms of vocal learning in zebra finches. Curr. Biol. *29*, 631–636.e3.

51. Jürgens, U. (1994). The role of the periaqueductal grey in vocal behaviour. Behav. Brain Res. *62*, 107–117.

52. Vicario, D.S., and Simpson, H.B. (1995). Electrical stimulation in forebrain nuclei elicits learned vocal patterns in songbirds. J. Neurophysiol. *73*, 2602–2607.

53. Karten, H.J., Brzozowska-Prechtl, A., Lovell, P.V., Tang, D.D., Mello, C.V., Wang, H., and Mitra, P.P. (2013). Digital atlas of the zebra finch (*Taeniopygia guttata*) brain: a high-resolution photo atlas. J. Comp. Neurol. *521*, 3702–3715.

54. Mello, C.V., Kaser, T., Buckner, A.A., Wirthlin, M., and Lovell, P.V. (2019). Molecular architecture of the zebra finch arcopallium. J. Comp. Neurol. *527*, 2512–2556.

55. Sen, S., Parishar, P., Pundir, A.S., Reiner, A., and Iyengar, S. (2019). The expression of tyrosine hydroxylase and DARPP-32 in the house crow (*Corvus splendens*) brain. J. Comp. Neurol. *527*, 1801–1836.

56. Kersten, Y., Friedrich-Müller, B., and Nieder, A. (2022). A brain atlas of the carrion crow (Corvus corone). J. Comp. Neurol. *530*, 3011–3038.

57. von Eugen, K., Tabrik, S., Güntürkün, O., and Ströckens, F. (2020). A comparative analysis of the dopaminergic innervation of the executive caudal nidopallium in pigeon, chicken, zebra finch, and carrion crow. J. Comp. Neurol. *528*, 2929–2955.

58. Tokarev, K., Tiunova, A., Scharff, C., and Anokhin, K. (2011). Food for song: expression of c-Fos and ZENK in the zebra finch song nuclei during food aversion learning. PLoS One *6*, e21157.

59. Chakraborty, M., and Jarvis, E.D. (2015). Brain evolution by brain pathway duplication. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20150056.

60. Hyland Bruno, J., Jarvis, E.D., Liberman, M., and Tchernichovski, O. (2021). Birdsong learning and culture: analogies with human spoken language. Annu. Rev. Linguist. *7*, 449–472.

61. Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M., and Lemmon, A.R. (2015). A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature *526*, 569–573.

62. Hage, S.R., and Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. Nat. Commun. *4*, 2409.

63. Hage, S.R., Gavrilov, N., and Nieder, A. (2013). Cognitive control of distinct vocalizations in rhesus monkeys. J. Cogn. Neurosci. *25*, 1692–1701.

64. Hage, S.R., Gavrilov, N., and Nieder, A. (2016). Developmental changes of cognitive vocal control in monkeys. J. Exp. Biol. *219*, 1744–1749.

65. Jürgens, U. (2002). Neural pathways underlying vocal control. Neurosci. Biobehav. Rev. *26*, 235–258.

66. Jürgens, U., and Ploog, D. (1970). Cerebral representation of vocalization in the squirrel monkey. Exp. Brain Res. *10*, 532–554.

67. Jürgens, U. (1976). Reinforcing concomitants of electrically elicited vocalizations. Exp. Brain Res. *26*, 203–214.

68. Bruce, C.J., Goldberg, M.E., Bushnell, M.C., and Stanton, G.B. (1985). Primate frontal eye fields. II. Physiological and anatomical correlates of electrically evoked eye movements. J. Neurophysiol. *54*, 714–734.

69. West, R.A., and Larson, C.R. (1995). Neurons of the anterior mesial cortex related to faciovocal activity in the awake monkey. J. Neurophysiol. *74*, 1856–1869.

70. Coudé, G., Ferrari, P.F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., Monti, F., Rozzi, S., and Fogassi, L. (2011). Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. PLoS One *6*, e26822.

71. Gavrilov, N., Hage, S.R., and Nieder, A. (2017). Functional specialization of the primate frontal lobe during cognitive control of vocalizations. Cell Rep. *21*, 2393–2406.

72. Gavrilov, N., and Nieder, A. (2021). Distinct neural networks for the volitional control of vocal and manual actions in the monkey homologue of Broca's area. Elife *10*, e62797.

73. Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J. Neurophysiol. *86*, 1916–1936.

74. Coe, B., Tomihara, K., Matsuzawa, M., and Hikosaka, O. (2002). Visual and anticipatory bias in three cortical eye fields of the monkey during an adaptive decision-making task. J. Neurosci. *22*, 5081–5090.

75. Williams, Z.M., Elfar, J.C., Eskandar, E.N., Toth, L.J., and Assad, J.A. (2003). Parietal activity and the perceived direction of ambiguous apparent motion. Nat. Neurosci. *6*, 616–623.

76. Carnevale, F., de Lafuente, V., Romo, R., and Parga, N. (2012). Internal signal correlates neural populations and biases perceptual decision reports. Proc. Natl. Acad. Sci. USA. *109*, 18938–18943.

77. Hanks, T.D., Mazurek, M.E., Kiani, R., Hopp, E., and Shadlen, M.N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. J. Neurosci. *31*, 6339–6352.

78. Dorris, M.C., Paré, M., and Munoz, D.P. (2000). Immediate neural plasticity shapes motor performance. J. Neurosci. *20*, 1–5.

79. Gold, J.I., Law, C.T., Connolly, P., and Bennur, S. (2008). The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. J. Neurophysiol. *100*, 2653–2668.

80. Akrami, A., Kopec, C.D., Diamond, M.E., and Brody, C.D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature *554*, 368–372.

81. Hermoso-Mendizabal, A., Hyafil, A., Rueda-Orozco, P.E., Jaramillo, S., Robbe, D., and de la Rocha, J. (2020). Response outcomes gate the impact of expectations on perceptual decisions. Nat. Commun. *11*, 1057.

82. Mochol, G., Kiani, R., and Moreno-Bote, R. (2021). Prefrontal cortex represents heuristics that shape choice bias and its integration into future behavior. Curr. Biol. *31*, 1234–1244.e6.

83. Estebanez, L., Hoffmann, D., Voigt, B.C., and Poulet, J.F.A. (2017). Parvalbumin-expressing GABAergic neurons in primary motor cortex signal reaching. Cell Rep. *20*, 308–318.

84. Melzer, S., Gil, M., Koser, D.E., Michael, M., Huang, K.W., and Monyer, H. (2017). Distinct corticostriatal GABAergic neurons modulate striatal output neurons and motor activity. Cell Rep. *19*, 1045–1055.

85. Swanson, O.K., and Maffei, A. (2019). From hiring to firing: activation of inhibitory neurons and their recruitment in behavior. Front. Mol. Neurosci. *12*, 168.

86. Hautus, M.J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d. Behav. Res. Methods Instrum. Comput. *27*, 46–51.

87. Szűcs, A. (1998). Applications of the spike density function in analysis of neuronal firing patterns. J. Neurosci. Methods *81*, 159–167.

88. Wobbrock, J.O., Findlater, L., Gergle, D., and Higgins, J.J. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In Conference: Proceedings of the International Conference on Human Factors in Computing Systems, CHI 2011, pp. 143–146.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Experimental models: Organisms/strains | | |
| Corvus corone | University of Tübingen, Institute of Neurobiology | bird 1, bird 2 |
| Software and algorithms | | |
| NIMH Cortex | National Institute of Mental Health | c595; https://www.nimh.nih.gov/labs-at-nimh/research-areas/clinics-and-labs/ln/shn/software-projects.shtml |
| MAP Data Acquisition System | Plexon | https://plexon.com/ |
| MATLAB R2017a | MathWorks | https://www.mathworks.com |
| Other | | |
| Dental Cement | Heraeus | Paladur, ISO 20795, CE 0197 |
| Microdrives | Animal Physiology Unit | Custom fabrication |
| Electrodes | Alpha Omega LTD | Cat.#: 366-130620-00 www.alphaomega-eng.com |

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Andreas Nieder (andreas.nieder@uni-tuebingen.de).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

All data reported in this paper will be shared by the lead contact upon request. This paper does not report original code. Any additional information required to re-analyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND SUBJECT DETAILS

#### Subjects

Data were collected from two male carrion crows (*Corvus corone corone*), aged 24 and 30 months (late adolescence/early adulthood) at the time of data collection, respectively. They earned food during the daily recordings, but, if necessary, further food was provided after the recording session. Body weight was measured daily. Water was provided ad libitum in the aviary and during testing. All procedures were approved by the local authorities (Regierungspräsidium Tübingen), and conducted in accordance with German and European law and the Guidelines of the National Institutes of Health for the Care and Use of Laboratory Animals.

### METHOD DETAILS

#### Set-up

The crows were trained and tested in a darkened and sound-attenuated operant conditioning chamber. Crows were perched in front of a touch screen monitor (3M Microtouch, 15″, 60-Hz refresh rate) that presented the task and registered their response. The CORTEX program (available at ftp://ftp.cnl.salk.edu/pub/cortex/, National Institute of Mental Health, MD, USA) controlled stimulus presentation. Vocalizations were recorded using a Sennheiser MKE 600 microphone, with a sampling rate of 40 kHz for offline analysis. The same microphone was used for the online detection of vocalizations as was used for the recording of the vocalizations for later analysis. Rewards (bird food pellets or mealworms) for correct trials were delivered with an automated feeder below the screen. In addition, crows received auditory feedback (a brief sound of a bell) for correct responses. An infrared light barrier, in combination

with a reflector foil attached to the crows' head, was used to ensure that the crow was positioned in front and facing the screen during the task.

### Behavioral protocol

Crows were trained on a detection task in which they had to vocalize at the screen in response to the detection of a visual go-cue to receive a reward. In this computerized go/nogo detection task (see Figure 1A), the crows earned rewards by vocalizing in response to a specific visual cue (''go cue''). The crows had to position their head in front of the monitor to close the infrared light barrier to start a trial. After a variable waiting period (1–5 s) during which a white square was shown, a go-cue (blue or red square) instructed the crows to vocalize, which was rewarded by food. This go-cue was presented for 3 s. Only vocalizations within these 3 s were rewarded and counted as a ''hit''. If the crows failed to vocalize to the presentation of the go-cue, the trial was counted as a ''miss''. In order to ensure that crows did not time their vocalization, in 11% of trials the go-cue did not appear after the waiting period (catch trials), and crows had to wait in silence until the cue disappeared. Vocalizations during catch trials, and during the waiting period, were defined as ''false alarms'', which were followed by a short time-out delaying the start of the next trial. Neither correct rejections nor misses were rewarded or punished. Vocalizations were detected automatically by a custom-built MATLAB program.

### Surgery and neurophysiological recordings

All surgeries were performed while the animals were under general anesthesia. They were anesthetized with a ketamine (50mg/kg body weight) and Rompun mixture (5 mg/kg xylazine) initially and, if necessary supplemented. After the surgery, the crows received analgesics (Butorphanol, 1 mg/kg). The head was placed in a customized stereotaxic holder. Using stereotaxic coordinates (center of craniotomy: anterior-posterior +5 mm relative to inter-aural (ear bars) as zero; medial-lateral 13 mm relative to midline), we chronically implanted two micro-drives with four electrodes each (spaced apart approximately 0.5 mm) in the right hemisphere targeting the medial part of the NCL (Kersten et al., 2022). Glass-coated tungsten microelectrodes with 2 MΩ impedance (Alpha Omega Co.) were used. The location of the recording site has been used previously, and has been histologically confirmed (Veit et al., 2014). In addition, a micro-connector for the head stage, and a retainer of the reflector foil was implanted on the scull. Extra-cellular single-cell activity was recorded in synchrony with task performance using the Plexon Multi-Acquisition System (for details, see Ditz et al.[33]). Neural signals as well as the vocalizations were digitized at a sampling rate of 40 kHz and stored to the PC running the Plexon system. Plexon's Offline Sorter was used to manually offline sort spikes into single-unit waveforms by applying mainly principal component analysis. On average, 1.5 neurons per active recording site were detected based on offline sorting.

The crows participated in daily recording sessions. At the beginning of each session, the electrodes were manually slightly advanced until a neuronal signal was detected on at least one of the electrodes. Single cells were separated offline (Plexon Offline Sorter); hence, during recording, the signals were not selected for their task involvement.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Behavioral analysis

To ensure precise timing and avoid false positives, the vocalizations were classified and validated manually with the aid of a custom-written MATLAB program. Analyses were performed in MATLAB R2015a and R. Vocalizations during go-trials were defined as ''hits'', vocalizations during catch-trials as ''false alarms'', respectively. A failure to vocalize during a go-trial was defined as a miss. Sensitivity values $d'$ derived from signal detection theory were calculated by subtracting z-scores (normal deviates) of median ''hit'' rates from z-scores of median ''false alarm'' rates ($d' = z$(hit rate) – $z$(false alarm rate)). To ensure that $d'$ was above the threshold value of 1.8, a Fisher Pitman permutation test was calculated for both crows separately. Crows completed on average 427 ± 104 and 253 ± 78 trials, respectively, per session. Because of false-alarm rates of 0 (see Results section), $d'$-estimates were corrected by a log-linear approach where 0.5 is added to the frequency of false alarms in each cell of the contingency table.[86]

### Neuronal analysis

#### *Vocalization-correlated neurons*

Neurons were recorded without pre-selection for response selectivity and sorted offline blind to task involvement. Neurons were included in further analyses if they showed a discharge rate of >0.5 Hz in the 2000 ms prior to response onset and could be recorded for at least 3 trials for both hit and miss trials. Average trial repetition was 255 (hits) and 11 (misses) for crow 1, and 153 (hits) and 10 (misses) for crow 2, respectively. Baseline activity was calculated in the 500 ms-period prior to the onset of trial instruction.

Vocalization-correlated neurons were detected by comparing neuronal activity (firing rates) between trials in which the crow vocalized to the presentation of the go-cue (hit trials) and trials in which the crow failed to vocalize to the presentation of the same cue (miss trials). This comparison was focused on a window of 1000 ms prior to the actual vocal response for ''hit'' trials, or to the average of the response latency in hit trials for ''miss'' trials. Because homogeneity of variances was not given for all cells (defined as $p < 0.05$ in a Bartlett's test and ratio of variances >4), data were analyzed using a non-parametric Wilcoxon rank-sum test. To isolate the neural correlate of volitional vocalizations, we tested whether there was an effect of the factor trial type on the average firing rate of each single unit. Single units for which this comparison was significant were considered to be vocalization-correlated. Alpha level was set at $\alpha = 0.05$.

For each single cell included in the analysis, a spike density function was generated. To do so, for each trial spikes were counted in 10 ms bins. These spike bins were convoluted with a normalized Gaussian function ($\sigma$ = 50 ms) cut at ± 300 ms around the peak. Finally, trials within the respective conditions were averaged.[87]

### Baseline fluctuations

To explore the potential impact of baseline firing rate fluctuations on hits or misses (the ''trial types''), we examined the neurons' firing rates 500 ms prior to the onset of the go-cue (i.e., when only the waiting cue was presented on the screen). This comparison was done separately for different 'types' of vocalization-correlated neurons, i.e., neurons that increased or decreased firing rates, respectively, after cue-onset but before vocal onset. To analyze baseline firing rates, we first used a repeated measure ANOVA for aligned rank transformed data[88] (factors ''trial type'' and ''type of neuron'') and subsequently a Wilcoxon signed rank test for the pairwise comparison (Bonferroni correction to control the family-wise error rate, $\alpha$ = 0.025).

### Task-unrelated vocalizations

We next compared volitional vocalization and vocalizations that the crows emitted in between trials and in trial breaks, e.g. when the crows saw the trainer (henceforth task-unrelated vocalizations). For comparison of neuronal preparatory activity before volitional versus task-unrelated vocalizations, we paired firing rates (of a given neuron) that were recorded when the acoustic parameters of volitional and task-unrelated vocalizations were best matched. We first characterized each (volitional and task-unrelated) vocalization by three vocalization parameters: duration, maximum frequency, and Wiener entropy (defined as natural logarithm of the quotient of the geometrical mean and the arithmetic mean of the power spectrum of the vocalization). We then determined which volitional vocalization in a session was closest (in parameter space) to a given task-unrelated vocalization for all three acoustic parameters by finding, for each task-unrelated vocalization, the volitional vocalization with the minimum Euclidian distance in the 3-dimensional parameter space. This volitional vocalization was then defined as matched to the task-unrelated vocalization. Finally, we compared the average preparatory firing rates (recorded during a minimum of three task-unrelated vocalizations) elicited to these matched volitional and task-unrelated vocalizations with each neuron as a data point in the paired distributions.

### Temporal dynamics

To investigate the temporal dynamics of vocalization-correlated activity prior to vocalization onset, we ran a sliding Wilcoxon rank-sum test on the vocalization specific neurons. The Wilcoxon rank-sum test was computed within 100 ms windows that were slid in 20 ms steps. That is, in each time window, we compared firing rates between trials with a vocalization (hit trials) and those without (miss trials). For significant time windows ($p < 0.05$), the effect size was calculated as area under the receiver operating characteristic curve (AUC) statistic as a standardized measure of the effect of trial type on average firing rate during these time windows. The latency of a cell was defined as the first time window in which there was a significant difference between firing rates for hit and miss trials.

### Neuronal population analyses

We analyzed activity in the entire population of recorded NCL neurons, irrespective of response preferences. We calculated for each trial the firing rate in 10 ms bins within a time window of −3000 ms (3000 ms before vocal onset) until 1000 ms after vocal onset. A spike density curve was then calculated by convoluting each bin with a Gaussian kernel with a standard deviation of 100 ms, cut at −300 to +300 ms. Cells with at least 10 hit and 10 miss trials were included in the analysis. For each trial type (hit/miss), 10 trials were randomly selected from the pool of all trials. This procedure generated a 20 × 58 × 401 sized data matrix, with 2x10 trials, 58 neurons and 401 time bins. This matrix was rearranged into a 8020x58 matrix. The data were then z-scored across the first dimension and a principal component analysis conducted. The mean across trials within each condition was computed and the Euclidian distance between both conditions was derived based on these means. For the classifier, in a leave-one-out procedure, for each trial the Euclidian distance to the mean across trials for each condition (without the respective trial) was computed. A trial was then classified as the condition to which it had the shorter distance. Performance was defined as the percentage of correct classifications. This procedure, starting with the random selection of trials, was repeated 10 times and the values averaged. To test whether classification behavior significantly exceeded chance behavior, we ran a randomization statistic. Here, the selected trials were randomly assigned to one of the two conditions and then the same procedure was applied. This procedure was repeated 1000 times and classification was defined as significant if it exceeded the 97.fifth percentile of the randomization distribution.