

Audio–vocal interactions during vocal communication in squirrel monkeys and their neurobiological implications

Steffen R. Hage

Received: 17 September 2012/Revised: 4 March 2013/Accepted: 5 March 2013/Published online: 21 March 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Several strategies have evolved in the vertebrate lineage to facilitate signal transmission in vocal communication. Here, I present a mechanism to facilitate signal transmission in a group of communicating common squirrel monkeys (*Saimiri sciureus sciureus*). Vocal onsets of a conspecific affect call initiation in all other members of the group in less than 100 ms. The probability of vocal onsets in a range of 100 ms after the beginning of a vocalization of another monkey was significantly decreased compared to the mean probability of call onsets. Additionally, the probability for vocal onsets of conspecifics was significantly increased just a few hundreds of milliseconds after call onset of others. These behavioral data suggest neural mechanisms that suppress vocal output just after the onset of environmental noise, such as vocalizations of conspecifics, and increase the probability of call initiation of group mates shortly after. These findings add new audio–vocal behaviors to the known strategies that modulate signal transmission in vocal communication. The present study will guide future neurobiological studies that explore how the observed audio–vocal behaviors are implemented in the monkey brain.

Keywords Acoustic communication · Audio–vocal integration · Brainstem mechanism · Inhibition · Vocal-motor control

Introduction

Successful signal transmission between sender and receiver faces the challenge of being subjected to masking noise during all communication processes in all animals (Brumm and Slabbekoorn 2005). For the auditory domain, animals have evolved several strategies to increase signal-to-noise ratio and, therefore, to facilitate the transmission of their acoustic signals. One example of such a mechanism is the Lombard effect, i.e., the involuntary rise in call amplitude and frequency in response to masking noise. This effect was first described in humans (Lombard 1911) and has since then also been found in birds (Cynx et al. 1998; Brumm and Todt 2002) and various mammals (Sinnott et al. 1975; Nonaka et al. 1997; Brumm et al. 2004; Hage et al. 2013; for review Brumm and Zollinger 2011). Another example for such a strategy is the avoidance of vocal output during the presence of environmental noise and, therefore, the timing of vocal initiation within silent gaps. This strategy has also been observed in frogs (Schwartz and Wells 1983), birds (Kobayasi and Okanoya 2003; Brumm 2006), and mammals (Roy et al. 2011). Animals must deal with another problem during vocal communication. They have to respond to the received vocalizations early enough to ensure that their own calls are recognized as coherent responses to the preceding vocal utterance. A previous behavioral study in black-headed squirrel monkeys (*Saimiri sciureus boliviensis*) showed that response latencies can be as short as 0.1–0.5 s during antiphonal calling, that is, in the reciprocal exchange of

Electronic supplementary material The online version of this article (doi:10.1007/s00359-013-0810-1) contains supplementary material, which is available to authorized users.

S. R. Hage (✉)
Animal Physiology, Institute of Neurobiology,
University of Tübingen, 72076 Tübingen, Germany
e-mail: steffen.hage@uni-tuebingen.de

S. R. Hage
Department of Neurobiology,
German Primate Center,
37077 Göttingen, Germany

contact calls between conspecifics (Masataka and Biben 1987). A recent playback experiment in marmoset monkeys demonstrated that response latencies are important in antiphonal calling behavior (Miller and Wang 2006). The probability to respond to a vocal utterance of a conspecific significantly decreased when this call was uttered with a latency of more than 9 s to a preceding vocalization of the receiving marmoset monkey (Miller et al. 2009). Therefore, especially in large groups of animals, strategies in vocal communication should consider both to avoid masking effects of environmental noise, i.e., vocalizations of conspecifics, and to leave the ability to communicate with adequate latency intact. In the present study, I analyzed the distribution of call onsets within a group of highly vocal squirrel monkeys (Winter et al. 1966; Baldwin 1968) and hypothesized that several mechanisms have evolved that facilitate signal transmission during vocal communication within this species. Based on these findings I suggest potential brain mechanisms underlying the observed vocal behavior.

Materials and methods

The study was carried out with ten male common squirrel monkeys (*Saimiri sciureus sciureus*). One monkey served as the focal animal and the other animals were used as its vocal partners. Surgical and experimental procedures were identical to those published previously (Hage and Jürgens 2006; Hage et al. 2006) and will only be summarized briefly. A platform was surgically implanted on the skull of the focal animal. Then, a piezo-ceramic skull vibration sensor was placed on the platform. This sensor served to distinguish the vocalizations of the experimental animal from those of the other animals of the group by comparing the signals picked up from the room microphone (ME64+K6, Sennheiser, Wedemark, Germany) and those originating from the skull vibration sensor. The signal of the sensor was fed into an amplifying transmitter circuit running with a 3-V battery and sent out with a carrier frequency between 100 and 150 MHz (for details on the electronic circuitry see Grohrock et al. 1997). The transmitter signal was picked up by an antenna within the animals' cage. The demodulated telemetric signal and the signal of the microphone placed in the animal room were sent to a personal computer via an A/D interface (Micro 1401 mkII, CED, Cambridge, UK) for data storage and offline analysis. Additionally, the telemetric signal was stored on a video recorder together with the room microphone signals and observational video camera recordings of the monkeys (for details on the telemetry setup, see Jürgens and Hage 2006).

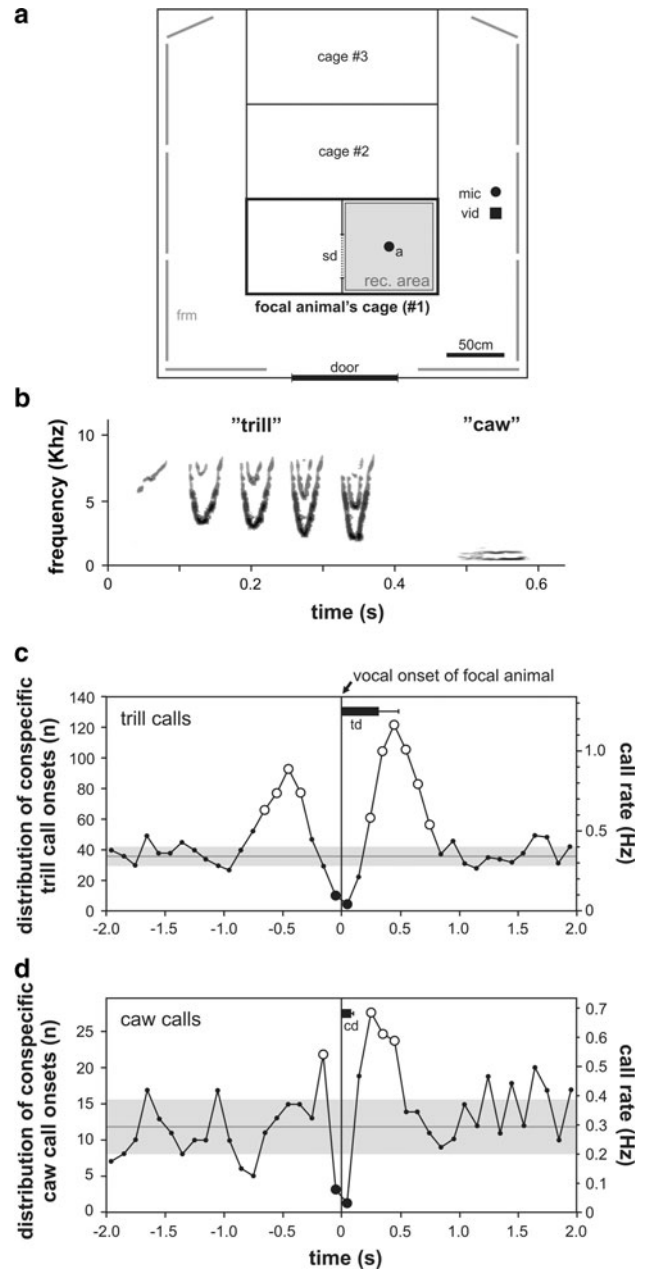
The focal animal was housed with two other animals in a cage of $2.4 (H) \times 0.8 (W) \times 1.6 (D) \text{ m}^3$. During

recording sessions, the animals were placed in one partition of their home cage [$2.4 (H) \times 0.8 (W) \times 0.8 (D) \text{ m}^3$] to ensure proper signal transmission by keeping the distance of the focal animal more or less constant to the antenna (which was placed in the middle of the cage; Fig. 1a). Two additional groups of three and four squirrel monkeys each were held in the same room in cages of $2.4 (H) \times 0.8 (W) \times 1.6 (D) \text{ m}^3$. Both groups had visual, acoustic as well as restricted tactile (only group #2) contact with the focal animal's group. The room was lined with foam rubber mats to reduce acoustic reflections (Fig. 1a) resulting in ambient background noise levels of 55–60 dB SPL. Vocal recording sessions lasted 10–15 min and were performed twice per day during feeding time, since frequency and diversity of vocalizations was the highest during that time. Before each session, the experimental animal was caught and placed in a monkey chair. The battery was exchanged, if necessary, and the animal was brought back to its home cage and recording started.

For data analysis, I focused on brief time windows of 10 s around vocal onsets of the focal animal, since I was particularly interested in short-term effects of call production. Trill and caw vocalizations were usually uttered during different behavioral contexts (Winter et al. 1966; Jürgens 1979) and were rarely uttered one after another. Trill vocalizations, highly frequency-modulated call types with a repetitive, rhythmic character and fundamental frequencies from 2 to 8 kHz, were typically uttered in response to food presentation (Winter et al. 1966). Caw calls, low-pitched calls with fundamental frequencies below 0.5 kHz, were vocalized during interactions between the experimenter and the monkey group. Therefore, I was able to investigate the vocal behavior of trill and caw calls separately. Rare cases where one call type was uttered within series of other call types were omitted from data analysis. Spontaneous call rates of conspecifics for both trill and caw calls were determined by calculating the averaged call rate of all conspecifics within 1–5 s before and after call onset of the focal animal. Vocal onsets of the partner animals were determined according to the onsets of the focal animal and grouped in 100 ms time bins. With this approach it was possible to determine whether vocal onsets of conspecifics occurred randomly or whether they were correlated with the vocal onset of the focal animal. If the latter case was true, the probability of vocal onsets of group mates would be elevated or reduced in specific time bins. Vocal responses of conspecifics to other conspecifics were not related to the vocal onset of the focal animal and, therefore, were randomly distributed around the vocal onset of the focal animal.

A one-sample Chi-square test was used to reveal whether probabilities of vocal onsets of partner animals were evenly distributed around the vocal onsets of the focal

Fig. 1 Vocal communication in squirrel monkeys. **a** Schematic drawing of the experimental setup (top view). The focal animal and its group mates were kept in the right section of their home cage (rec. area) during recording sessions. The two other groups were housed in cage #2 (three animals) and #3 (four animals); *a* antenna, *frm* foam rubber mats, *mic* microphone, *sd* sliding door, *vid* video camera for visual monitoring. **b** Example of spontaneously uttered trill and caw vocalizations of a squirrel monkey. Spectrograms as recorded with the bone vibration sensor. Intensity is represented by gray level. Probabilities of vocal onsets of group mates show significant non-uniform distributions for both trill (**c**) (one-sample Chi-square test, $p < 0.01$) and caw calls (**d**) (one-sample Chi-square test, $p < 0.001$) around the vocal onset of the focal animal (left ordinate). Gray horizontal lines indicate the averaged call rate of the conspecifics with standard deviations (shaded areas, right ordinate). Time bins (bin width 100 ms) with elevated or reduced numbers of vocal onsets of conspecifics compared to the averaged call rate are marked with open and filled circles, respectively (post hoc single linkage clustering analysis). Black horizontal bars indicate mean trill (*td*) and mean caw duration (*cd*), respectively (both with standard deviation)



animal or not. When significant differences were found, a post hoc clustering analysis (single linkage; proximity matrix with squared Euclidean distance) was conducted to identify significant time bins in the data set. I used the single linkage method (“nearest neighbor”), since it is the most reliable method to test for outliers in a data set, i.e., in this case, time bins that show an elevated or reduced number of vocal onsets compared to the number in the other time bins (Backhaus et al. 2011). The analyzed data sets were interval-scaled. Therefore, the squared Euclidean distance method was chosen as proximity matrix. All tests were made with SPSS 21 (IBM Corporation, USA). Differences in distributions were considered significant if the probability of error was $< 5\%$.

Results

I recorded 3,982 trill vocalizations of nine monkeys uttered within a 10-s time window (5 s prior to and 5 s after vocal onset of the focal animal). These vocalizations occurred around 1,043 spontaneously uttered trill calls of the focal animal. 1,193 caw vocalizations of the partner animals were uttered around 397 caw calls of the focal animal. Figure 1b shows an example of both spontaneously uttered trill and caw vocalizations, as recorded with the bone vibration sensor. Spontaneous call rates of conspecifics were observed with 0.34 ± 0.06 Hz for trill calls, and 0.29 ± 0.09 Hz for caw vocalizations. The timing of vocal onsets of group mates show significant non-uniform distributions around the vocal onsets of the focal animal in both trill and caw vocalizations [one-sample Chi-square test, $p < 0.01$ (trill), $p < 0.001$ (caw); Fig. 1c, d]. Group mates uttered trill vocalizations with higher probability 200–800 ms after vocal onsets of the focal animal

compared with the mean probability of vocal onsets of group mates (post hoc single linkage clustering analysis). Mean duration of trill vocalizations of the focal animal was 324.4 ± 162.2 ms. This indicates that conspecifics started to respond to trill calls of the focal animal already prior to or immediately after call offset. In addition, the probability of vocal onsets of group mates was higher about 300–700 ms prior to vocal onsets of the focal animal indicating that the focal animal responded to trills of its group mates with similar latency as vice versa (post hoc single linkage clustering analysis; Fig. 1c; Supplementary Fig. 1a, time bins marked with open circles). Similar results were obtained during the utterance of caw calls,

where higher probabilities of vocal onsets of group mates were found 100–200 ms prior to and 200–500 ms after the onset of the focal animal (post hoc single linkage clustering analysis; Fig. 1d; Supplementary Fig. 1b). The mean duration of caw calls of the focal animal was 78.5 ± 25.6 ms.

An opposite effect was obtained just around the vocal onsets of the focal animal in both trill and caw vocalizations. In caw sequences, the probability for vocal onsets of group mates was decreased compared to the mean call rate indicating that the focal animal did not start a vocalization directly after the onset of a group mate vocalization (post hoc single linkage clustering analysis; Fig. 1c, d, time bins marked with filled circles). Similarly, group mates did not begin a vocal utterance immediately after vocal initiation of the focal animal.

Discussion

I recorded trill and caw vocalizations of a group of squirrel monkeys during feeding sessions to investigate their vocal behavior and determine how vocal onsets are distributed within this group over time. Vocal recordings in a group of ten squirrel monkeys often result in a 100 calls per minute (personal observation). In such a situation, it is nearly impossible to resolve whether and, if so, to which particular sender an individual receiver is responding to. Several studies have overcome this problem by testing time intervals between calls by caging animals separately during tests (e.g., Miller and Wang 2006; Miller et al. 2009). Such experimental designs can not only elucidate who responds to whom, but also, for example, whether specific animals respond more readily to group mates or mates from outside the cage (e.g., Miller and Wang 2006). However, such an approach represents a pure psychophysical study and eliminates the ethological context. Therefore, I decided to relate vocal onsets of several monkeys to one single monkey which was equipped with a bone vibration sensor, to examine whether group mates modulated their vocal output in response to the focal animal and/or vice versa. This approach did not allow to analyze every single vocal interaction within the colony or to determine who responds to whom in every single case. However, it gave me the unique possibility to investigate rich vocal behavior within a monkey colony with a focal animal.

The results of the present study show that the vocal behavior within the group was similar—differences in response latencies were within one bin size—with the focal monkey that served as the receiver of preceding vocalizations or as the sender for subsequent vocalizations to the group mates. The data set does not provide information on whether the response latencies observed would have been different if I had chosen another monkey of the colony for

the experiment. The fact, however, that conspecifics respond to the focal animal with similar latencies as the focal animal responds to its conspecifics indicates that the present findings do not only apply to the focal animal, but to the entire group observed in this study. Therefore, it seems that these short response latencies do not depend on communicational rules as they can be found, for example, in antiphonal calling behavior of marmosets (e.g., Miller and Wang 2006).

The present data show that during the utterance of both call types vocal onsets were unequally distributed. Probabilities of vocal onset were significantly increased already 100–200 ms after caw call onset and 200–300 ms after trill call onset, respectively. Vocal behavior with similar response latencies has been observed in a previous study on antiphonal calling in a group of black-headed squirrel monkeys (*S. sciureus boliviensis*), another squirrel monkey subspecies with a slightly different call repertoire (Masataka and Biben 1987). In the previous study, black-headed squirrel monkeys showed short vocal response latencies of 100–500 ms during antiphonal calling with chuck calls that are characterized by a single downward frequency sweep with fundamental frequencies from 2 to 8 kHz and durations between 30 and 80 ms (Winter et al. 1966). Short response latencies of 0.5 s or less in vocal behavior, however, do not appear to exist in all monkey species. Mean response latencies of vocal utterances during antiphonal calling behavior of marmosets, for example, are around 5 s (Miller and Wang 2006). Due to the experimental design and high vocal call rates around the call onsets of the focal animal, it could not be determined in the present study whether animals respond more readily to mates in the same cage as opposed to mates outside.

Most strikingly, I observed a significantly decreased probability of vocal onsets of conspecifics just around call onsets of the focal animal. Such a decrease in call probability has not been reported during antiphonal calling in black-headed squirrel monkeys which was most likely due to the low spontaneous chuck call rate that was observed in this previous study (Masataka and Biben 1987). Focusing on caw vocalizations, the present data suggest that squirrel monkeys vocalize during silent gaps to avoid masking of the emitted signal as described in several other vertebrates including non-human primates (Schwartz and Wells 1983; Kobayasi and Okanoya 2003; Brumm 2006; Roy et al. 2011). However, the obtained data set on trill vocalizations argue for a different interpretation. Similar to caw vocalizations, the probability of trill call onsets of conspecifics was decreased just 100 ms before and after the onset of the trill call of the focal animal, i.e., there was substantial overlap between the trill call of the sender and the response of the receiver. Trill vocalizations of the

receiver, however, still persisted after trill call offset of the sender due to the delay of call onsets. Since trill vocalizations are repetitive and rhythmic in structure, small persisting fractions might be sufficient to be recognized by the sender as a proper trill vocalization (Winter et al. 1966). Therefore, it seems that squirrel monkeys combine two important strategies in vocal communication. On the one hand, they avoid masking effects of vocalizations of conspecifics to a specific amount which is one of the main issues in vocal communication (Brumm and Slabbekoorn 2005). On the other hand, they possess the ability to communicate with short response latency to a sender, which might be of considerable importance especially in large monkey colonies.

The underlying neuronal networks that are responsible for such fast modulations of vocal output are yet unknown. However, response latencies as observed in the present study suggest that the vocal-motor network is modulated on brainstem level. It is known from several studies (for review e.g., Jürgens 2002; Jürgens and Hage 2007) that the periaqueductal gray in the midbrain and the vocal pattern generator in the ventrolateral pontine brainstem are crucial components in the vocal-motor network. Both structures are directly involved in triggering the onset of vocal utterances. Single-unit studies have revealed short mean lead times of 100 ms prior to vocal output in the ventrolateral reticular formation of monkeys (Hage and Jürgens 2006). In contrast, single neurons in the periaqueductal gray of monkeys changed their firing rates already around 400 ms prior to vocal output (Larson and Kistler 1984). Comparing the lead times of these brain structures with the observed latencies for decreased vocal output after the vocal onset of a conspecific suggests a modulation of the vocal-motor network by auditory structures not at the level of the periaqueductal gray, but instead on pontine level. Several structures have been identified within the pontine brainstem of monkeys and other mammals which seem to be involved in audio–vocal integration, such as the ventrolateral reticular formation (Hage et al. 2006), the ventral nucleus of the lateral lemniscus (Suga and Schlegel 1972; Suga and Shimozawa 1974; Hage et al. 2006) and the paralemniscal area (Metzner 1996). Further studies will have to show if and how these structures are involved in the adjustments observed in the present study.

In conclusion, vocal onsets of single individuals in a group of squirrel monkeys are highly dependent on each other. Vocal utterances inhibit call initiation just after call onset, while they increase the probability of vocalizations shortly after that inhibitory phase. In combination with published data on vocal-motor control mechanisms, these findings suggest brainstem-based audio–vocal integration processes that are involved in such short response latencies during intraspecific vocal behavior of squirrel monkeys.

Acknowledgments I thank Uwe Jürgens and two anonymous referees for their helpful comments on this manuscript. Furthermore, I thank Kurt Hammerschmidt for his helpful advice on statistical methods and Simon N. Jacob for proof reading. All research described here was approved by the Animals Ethics Committee of the District Government Braunschweig, Lower Saxony, Germany. The experiments conformed to the NIH guidelines on the ethical use of animals.

References

- Backhaus K, Erichson B, Plinke W, Weiber R (2011) *Multivariate Analysemethoden*. Springer, Berlin
- Baldwin JD (1968) The social behavior of adult male squirrel monkeys (*Saimiri sciureus*) in a semi-natural environment. *Folia Primatol* 9:281–314
- Brumm H (2006) Signaling through acoustic windows: nightingales avoid interspecific competition by short-term adjustment of song timing. *J Comp Physiol A* 192:1279–1285
- Brumm H, Slabbekoorn H (2005) Acoustic communication in noise. *Adv Stud Behav* 35:151–209
- Brumm H, Todt D (2002) Noise-dependent song amplitude regulation in a territorial songbird. *Anim Behav* 63:891–897
- Brumm H, Zollinger SA (2011) The evolution of the Lombard effect: 100 years of psychoacoustic research. *Behaviour* 148:1173–1198
- Brumm H, Voss K, Köllmer I, Todt D (2004) Acoustic communication in noise: regulation of call characteristics in a new world monkey. *J Exp Biol* 207:443–448
- Cynx J, Lewis R, Tavel B, Tse H (1998) Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata*. *Anim Behav* 56:107–113
- Grohrock P, Häusler U, Jürgens U (1997) Dual-channel telemetry system for recording vocalization-correlated neuronal activity in freely moving squirrel monkeys. *J Neurosci Methods* 76:7–13
- Hage SR, Jürgens U (2006) On the role of the pontine brainstem in vocal pattern generation. A telemetric single-unit recording study in the squirrel monkey. *J Neurosci* 26:7105–7115
- Hage SR, Jürgens U, Ehret G (2006) Audio–vocal interaction in the pontine brainstem during self-initiated vocalization in the squirrel monkey. *Eur J Neurosci* 23:3297–3308
- Hage SR, Jiang T, Berquist S, Feng J, Metzner W (2013) Ambient noise induces independent shifts in call frequency and amplitude within the Lombard effect in echolocating bats. *Proc Natl Acad Sci USA* 110:4063–4068
- Jürgens U (1979) Vocalization as an emotional indicator. A neuroethological study in the squirrel monkey. *Behaviour* 69:88–117
- Jürgens U (2002) Neural pathways underlying vocal control. *Neurosci Biobehav Rev* 26:235–258
- Jürgens U, Hage SR (2006) Telemetric recording of neuronal activity. *Methods* 38:195–201
- Jürgens U, Hage SR (2007) On the role of the reticular formation in vocal pattern generation. *Behav Brain Res* 182:308–314
- Kobayasi KI, Okanoya K (2003) Context-dependent song amplitude control in Bengalese finches. *NeuroReport* 14:521–524
- Larson CR, Kistler MK (1984) Periaqueductal gray neuronal activity associated with laryngeal EMG and vocalization in the awake monkey. *Neurosci Lett* 46:261–266
- Lombard E (1911) Le signe de l'élévation de la voix. *Ann Mal Oreille Larynx* 37:101–119
- Masataka N, Biben M (1987) Temporal rules regulating affiliative vocal exchanges of squirrel monkeys. *Behaviour* 101:311–319
- Metzner W (1996) Anatomical basis for audio–vocal integration in echolocating horseshoe bats. *J Comp Neurol* 368:252–269

- Miller CT, Wang X (2006) Sensory–motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. *J Comp Physiol A* 192:27–38
- Miller CT, Beck K, Meade B, Wang X (2009) Antiphonal call timing in marmosets is behaviorally significant: interactive playback experiments. *J Comp Physiol A* 195:783–789
- Nonaka S, Takahashi R, Enomoto K, Katada A, Unno T (1997) Lombard reflex during PAG-induced vocalization in decerebrate cats. *Neurosci Res* 29:283–289
- Roy S, Miller CT, Gottsch D, Wang X (2011) Vocal control by the common marmoset in the presence of interfering noise. *J Exp Biol* 214:3619–3629
- Schwartz JJ, Wells KD (1983) An experimental study of the acoustic interference between two species of Neotropical tree frogs. *Anim Behav* 31:891–897
- Sinnott JM, Stebbins WC, Moody DB (1975) Regulation of voice amplitude by the monkey. *J Acoust Soc Am* 58:412–414
- Suga N, Schlegel P (1972) Neural attenuation of responses to emitted sounds in echolocating bats. *Science* 177:82–84
- Suga N, Shimozawa T (1974) Site of neural attenuation of responses to self-vocalized sounds in echolocating bat. *Science* 183:1211–1213
- Winter P, Ploog D, Latta J (1966) Vocal repertoire of the squirrel monkey (*Saimiri sciureus*), its analysis and significance. *Exp Brain Res* 50:359–366

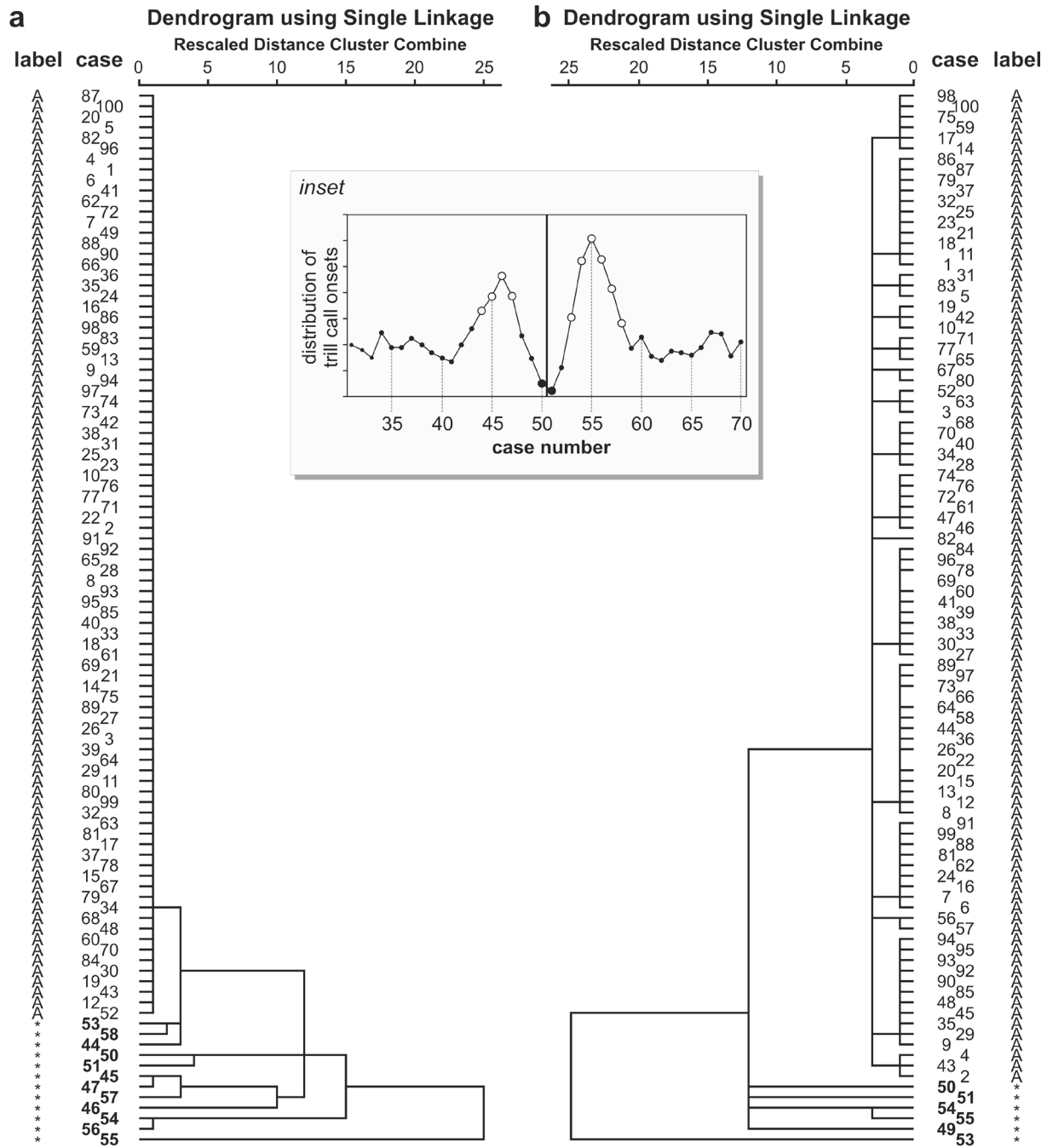


Figure S1 Dendrograms by using Single Linkage from the clustering analysis **a** In the dendrogram for the distribution of trill onsets, time bins 44-47, 50, 51 and 53-58 were found to be outliers (label *), while the remaining time bins clustered together (label A). **b** In the dendrogram of the distribution of caw onsets, time bins 49-51 and 53-55 were defined as outliers (label *), while the remaining time bins were combined into a single cluster (label A). The inset illustrates the assignment of cases to time bins for the distribution of probabilities of trill onsets, exemplarily. Note that the time window shows only a fraction (4s) of the entire analysis window (10s; see Material and Methods).