Current Biology 22, 2144–2148, November 20, 2012 ©2012 Elsevier Ltd All rights reserved http://dx.doi.org/10.1016/j.cub.2012.09.022

Report

An Asian Elephant Imitates Human Speech

Angela S. Stoeger,[1,6,*] Daniel Mietchen,[2,6] Sukhun Oh,[3] Shermin de Silva,[4,5] Christian T. Herbst,[1] Soowhan Kwon,[3] and W. Tecumseh Fitch[1,*]
[1]Department of Cognitive Biology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria
[2]EvoMRI Communications, Zwätzengasse 10, 07743 Jena, Germany
[3]Samsung Everland Zoo, 310 Jeondae-ri, Pogok-eup, Cheoin-gu, Yongin-si, Gyeonggi-do 449-715, South Korea
[4]Department of Fish, Wildlife, and Conservation Biology, Colorado State University, 1474 Campus Delivery, Fort Collins, CO 80523, USA
[5]Elephant Forest and Environment Conservation Trust, 215 A 3/7 Park Road, Colombo 5, Sri Lanka

Summary

Vocal imitation has convergently evolved in many species, allowing learning and cultural transmission of complex, conspecific sounds, as in birdsong [1, 2]. Scattered instances also exist of vocal imitation across species, including mockingbirds imitating other species or parrots and mynahs producing human speech [3, 4]. Here, we document a male Asian elephant (*Elephas maximus*) that imitates human speech, matching Korean formants and fundamental frequency in such detail that Korean native speakers can readily understand and transcribe the imitations. To create these very accurate imitations of speech formant frequencies, this elephant (named Koshik) places his trunk inside his mouth, modulating the shape of the vocal tract during controlled phonation. This represents a wholly novel method of vocal production and formant control in this or any other species. One hypothesized role for vocal imitation is to facilitate vocal recognition by heightening the similarity between related or socially affiliated individuals [1, 2]. The social circumstances under which Koshik's speech imitations developed suggest that one function of vocal learning might be to cement social bonds and, in unusual cases, social bonds across species.

Results and Discussion

Vocal learning, a crucial component of human speech, has evolved independently in several distantly related taxa, typically to allow the learning and cultural transmission of complex, conspecific calls [1, 2]. The learned songs of birds [5–8] and whales [9] are the best-known examples. Numerous instances of vocal imitation across species (sometimes termed "vocal mimicry") also exist, for example animals imitating human speech. Among birds, parrots and mynahs are talented imitators of the human voice [3, 4], but only a few convincing examples of speech imitation in nonhuman mammals are known. One documented case was Hoover, a harbor seal (*Phoca vitulina*) who could utter simple phrases in English after being raised by a Maine fisherman [10]. Another study documented that an adult male beluga (*Delphinapterus leucas*) imitated his name "Logosi" [11]. Anecdotal reports further suggest that a male Asian elephant (*Elephas maximus*) in a zoo in Kazakhstan might have been capable of producing speechlike utterances in Russian and Kazakh [12], but documentation is lacking.

Human speech imitation in animals requires a complex match between vocal perception and production to perceive, decode, and reproduce the speech signal. Despite considerable effort, several attempts to train apes to imitate human speech provide little support for ape vocal imitation abilities [13]. The inability of our nearest living relatives to imitate speech apparently stems from poor cortical-motor control of the larynx and the vocal tract [14–16]. Despite lacking certain morphological structures that humans use to articulate speech sounds (e.g., having a beak instead of lips), some animals can overcome morphological constraints that might seem to preclude production of human sounds, as long as neuronal circuitry specialized for perceiving and reproducing an acoustic signal is available.

Here, we analyze human speech imitation by a male Asian elephant named Koshik from the Everland Zoo in South Korea, augmenting and extending prior evidence of vocal imitation in elephants [17].

Speech Imitative Repertoire

Koshik's speech sound repertoire was said by his trainers to comprise six Korean words. We tested this hypothesis by analyzing transcriptions made by 16 Korean native speakers on 47 recordings of Koshik's utterances (see Table S1 available online). The subjects were not informed about the supposed spelling or meaning of the imitations. This analysis largely confirmed the trainers' claims, indicating that Koshik's speech imitations correspond to the following five words: "annyong" ("hello," Audio S1), "anja" ("sit down," Audio S2), "aniya" ("no"), "nuo" ("lie down," Audio S3), and "choah" ("good," Audio S4). Agreement was high for vowels and relatively poor for consonants: vowel transcription similarity was 67% overall, whereas consonant agreement only reached 21% (Table S1). For example, "choah" utterances (according to trainers) were mainly transcribed as "boah" ("look," 38%) or "moa" ("collect," 23%), but neither of these utterances was used toward Koshik. As a result, transcriptions provided exact spelling matches (in Korean) for only one sound ("annyong," "hello," for which the majority of respondents [56%] agreed) and three additional imitations for which considerable agreement could be documented ("aniya": 44%; "nuo": 31%; "anja": 15%). These results show that Koshik accurately imitates vowels, determined by formant frequency matching, but that consonant fidelity is relatively poor. Korean is not a tonal language like Chinese, in which changes in fundamental frequency are phonemic and change word meanings. Figure 1 contains spectrographic depictions of Koshik's speech imitation corresponding to the word "nuo," together with "nuo" produced by one of his trainers and a native Korean speaker unfamiliar with Koshik.

Comparison of the Elephant's Speech Imitation, Human Speech, and Natural Asian Elephant Calls

We applied discriminant function analysis (DFA) to compare structural characteristics of Koshik's speech imitations to

[6]These authors contributed equally to this work
*Correspondence: angela.stoeger-horwath@univie.ac.at (A.S.S.), tecumseh.fitch@univie.ac.at (W.T.F.)

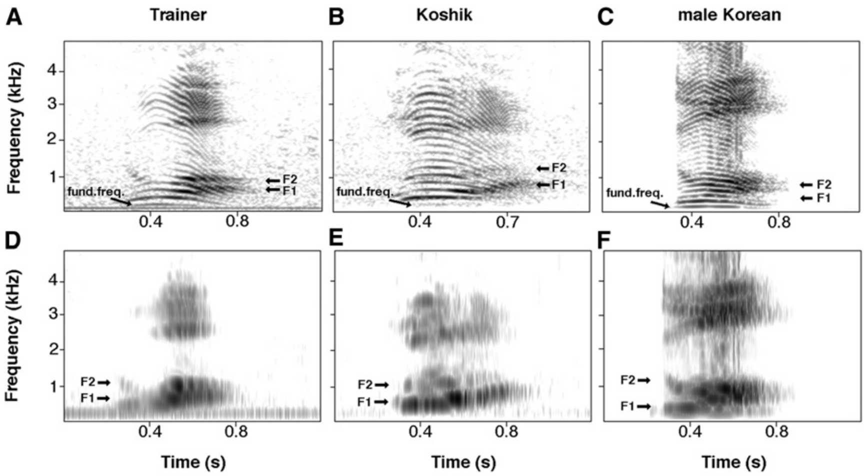

Figure 1. Spectral Comparison of the Speech Utterance "nuo"

Spectrograms exemplifying the speech utterance "nuo" of the trainer (A and D) compared to the elephant's (Koshik) imitation (B and E) and a 40-year-old male Korean native speaker (C and F) with no experience of Koshik's Korean output (recorded via a head set and thus with higher recording quality than the other two sound samples). (A–C) represent narrow band spectrograms of "nuo" and (D–F) give wide-band spectrograms of each "nuo" utterance, respectively. The fundamental frequency (fund. freq.) and the first and the second formant (F1 and F2) are indicated.

natural Asian elephant calls (using duration, minimum and maximum fundamental frequency, and the first formant/spectral peak frequency), finding that Koshik's imitations are very different from 187 calls of 22 Asian elephants of both genders and various ages recorded in five different zoos and in the Udawalawe National Park, Sri Lanka (Table S2). Instead, they cluster tightly with the human model utterances (Figure 2A), which were recorded from Koshik's trainers. Fundamental frequency is the most discriminating feature. Post hoc Bonferroni tests revealed no significant difference in minimum or maximum fundamental frequency between Koshik's imitations and the trainer's utterances, but showed significant differences relative to the natural Asian elephant calls (all $p < 0.001$) (Figure 2B).

Koshik's Speech Production and Formant Matching

Particularly during vowel production, Koshik's first two formants accurately match formant 1 and formant 2 of his trainers (Figure 3). Comparing means of the first and second formant with the corresponding human formant of the most commonly recorded vowels, "a," "o," and "u," revealed no significant difference between the elephant and the human models (Table S3). Koshik's precise imitation of the acoustic characteristics of his trainers is remarkable, given that the long vocal tract of an elephant would naturally produce much lower formant frequencies [19]. Koshik creates these accurate imitations of human formant frequencies by placing his trunk tip into his mouth (always from the right side; Figure 3A and Movie S1, Movie S2, and Movie S3) at the onset of phonation (about 0.3 ± 0.11 s before starting to vocalize, n = 50). During phonation, he raises the lower jaw while keeping the trunk inside the mouth, thus modulating the shape of his vocal tract. The elephant removes the trunk from the oral cavity about 0.4 ± 0.23 s (n = 50) after phonation. There is no considerable difference in the timing of trunk insertion and removal between the different imitations.

Not much is known about Asian elephant sound production in general. Presumably, low-frequency rumbles are produced via the same physiological production mechanism (passive vocal-fold vibration) as in human speech, as recently shown for African elephants [20]. Whether this is true for all call types, and whether particular elephant sounds are emitted nasally or orally, remains unknown. In any case, Koshik's use of his trunk to produce speech sounds is very unusual and has not been reported for wild Asian elephants [21, 22], nor for Koshik when he produces natural elephant calls. Three other Asian elephants have been described to whistle by pressing the trunk against the mouth [23]. Putting a body part, in Koshik's case the trunk, inside the mouth, thereby modulating the vocal tract in order to manipulate formants, is a wholly novel method of vocal production. Lacking X-ray images, we cannot be certain whether tongue movements are also involved in Koshik's speech imitations. But we do know that elephants lack a full oral sphincter, because the upper lip is fused with the nose to form the trunk. Lip rounding, a feature of vowels such as /u/, is thus, strictly speaking, impossible. Koshik's success at vowel imitation suggests that elephants are able to overcome morphological limitations by augmenting the oral vocal tract with their trunk: an evolutionarily novel and highly specialized appendage. The only vaguely reminiscent result we are aware of, outside of humans, concerns orangutans (*Pongo pygmaeus wurmbii*), who are reported to modulate sound spectra using their hands or leaves [24].

The results indicate that the elephant brain can transfer detailed information between auditory centers and the corresponding motor planning regions (including those controlling the trunk muscles), in addition to having the precise control over the larynx necessary to gate and modulate fundamental frequency. Our documentation of elephant vocal learning adds support to the "vocal learning and rhythmic synchronization hypothesis," since it has been recently suggested that Asian elephants may be capable of beat perception and synchronization (BPS) [25, 26]. This hypothesis signifies that entrainment might have evolved as a byproduct of selection for vocal imitation (BPS also requires information transfer between the auditory and motor systems) and, thus, that only vocal learning species should be capable of BPS [25, 26]. The alternative, that entrainment leads to vocal imitation, is rendered unlikely by the finding that, while all known entraining species are vocal learners, many vocal learners show no entrainment ability [26].

Although elephants living under human care may be heavily exposed to speech from birth on, they do not imitate speech on a regular basis. Thus, early intensive speech exposure does not seem adequate to initiate speech imitation in elephants (although it might be a required precondition), as long as they are embedded within an elephant social environment. Koshik was captive-born in 1990 and translocated to Everland in 1993, where two female Asian elephants accompanied him until he was five years old. From 1995 to 2002, Koshik was the only elephant in Everland. He was trained to physically obey several commands and was exposed to

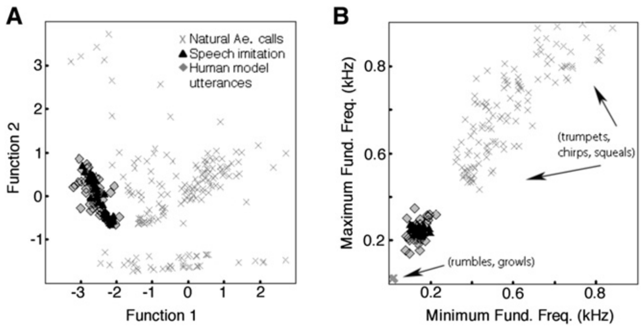

Figure 2. Comparison of Koshik's Imitations and Human Utterances with Natural Asian Elephant Calls

(A) Scatterplot representing function 1 and 2 of the DFA. None of Koshik's imitations was classified with the natural Asian elephant (Ae.) calls, whereas 50% were allocated to the human utterances. In turn, 58% of the human utterances (trainers) were allocated to Koshik's imitations. The strongest factor loading of the first function included the variables maximum (0.477) and minimum (0.441) fundamental frequency (% variance explained: 99.5%), with formant/spectral peak frequency (0.997) on the second function (% variance explained: 0.5%).

(B) Scatterplot showing the distribution of utterances based on the minimum and maximum fundamental frequency (fund. freq.).

human speech intensively by his trainers, veterinarians, guides, and tourists. In August 2004, his trainers first noticed that Koshik imitated speech. We cannot be certain whether Koshik started to produce speech sounds at 14 years of age (near the onset of Koshik's sexual maturity; his first musth period [27] occurred in March 2005) or whether earlier imitations went unrecognized by his trainers. However, the determining factors for speech imitation in Koshik may be social deprivation from conspecifics during an important period of bonding and development when humans were the only social contact available (this hypothesis may also hold for other known examples of speech imitation in mammals, Hoover the seal and the beluga Logosi, and also most talking birds [28–30]).

Together with previous examples documenting vocal production learning in African elephants [17], these new data extend the vocal learning ability to both surviving genera of the once-numerous Elephantidae [31]. What function or functions might vocal learning serve in elephants? In seals, baleen whales, and many passerine species, which vocalize or sing to attract mates and/or defend territories, vocal learning might facilitate the generation of more complex songs or calls and thus increase reproductive success via sexual selection [1, 2, 5, 7–9]. In elephants, little is known about the functional relevance of male calls, which males produce more frequently during musth periods [27]. Koshik, however, produced speech imitations throughout the year, not only when in musth.

Another potential role of vocal imitation might be to facilitate vocal recognition by heightening the similarity between related or socially affiliated individuals. Convergence of vocal signals as animals become associated has been reported for a wide range of birds and mammals [32]. This is particularly important in highly dynamic and flexible social systems [2, 33], in which multiple social groups overlap in space and do not hold spatially exclusive territories. Such systems are characteristic of many vocal learners, including elephants [34, 35] and even humans. The social circumstances under which Koshik's speech imitations developed suggest that one function of vocal learning might be to cement social bonds and, in unusual cases, social bonds across species.

Experimental Procedures

Acoustic Data Collection

Koshik was recorded in an outdoor enclosure (microphone-to-elephant distance of 10 to 25 m) of the Everland Zoo from October 3–8, 2010, on a daily basis for 25 hr in total. He was typically stimulated to vocalize: A trainer or the veterinarian said his name or "hello Koshik" ("Koshik annyong" in Korean) and continued talking to Koshik with words from his imitative repertoire until Koshik responded. This typically occurred within two to three utterances. Although Koshik's speechlike vocalizations often followed such "requests" from his trainers, he frequently uttered different words than those used immediately previously by his trainers. Koshik was usually rewarded after each imitation and, therefore, was not specifically trained to only reproduce the preceding utterance. Additionally, Koshik sometimes spontaneously produced speechlike utterances (mainly "choah" and "nuo"). These spontaneous vocalizations did not vary from the ones recorded during the interactions with the trainers. We recorded 320 imitative calls. We recorded the two trainers and the veterinarian simultaneously with the elephant, while they worked and interacted with him (80 words from trainer 1, 50 words from trainer 2, and 30 words from the veterinarian).

Acoustic recordings were made with two AKG 480 B microphones with condenser capsule CK 62 (frequency response: 8 Hz to 20 kHz ± 0.9 dB) and a Zoom 300 digital recorder. We used a Canon Legria FS200 Camcorder connected to an external Sennheiser K6 microphone with an ME 67 capsule for video recordings.

The various speech sounds (of Koshik and the humans) were annotated and extracted using the PRAAT 5.0.29 DSP package. In addition, we annotated and extracted time units of 0.15 s of the most common vowels (particularly "a," "o," and "u") within each imitative vocalization.

In order to compare Koshik's imitations of human speech with the natural vocal repertoire of *E. maximus*, we analyzed a large set of calls of free-ranging Asian elephants (n_{calls} = 3921, where caller and context were observed for 620 calls, including 11 males and 51 females) and Asian elephants recorded in zoos (n_{calls} = 302, of 6 male and 6 female Asian elephants). Recordings from free-ranging elephants in Udawalawe National Park (UWNP), Sri Lanka (May 2006 to December 2007), were made on a Fostex FR-2 field recorder (sampling rate 48 kHz) connected to a 12 V lead acid battery and an Earthworks QTC50 microphone (for details on the recording methods see [22]).

Recordings of the zoo animals were done at the Everland Zoo, the Leipzig Zoo (April 2011, 50 hr), the Walding Zoo (2007 to 2009, 60 hr), and the Heidelberg Zoo (April 2012, 20 hr), using the same equipment as for the recordings of Koshik. In the Emmen Zoo (2002, 30 hr), recordings were made using a Tascam DA-P1 DAT recorder (same microphones as above).

Elephant calls were annotated and extracted using the PRAAT v4.5.16. (Sri Lanka data) and PRAAT 5.0.29 DSP package (zoo data).

Questionnaire

A questionnaire with sounds recorded from Koshik was administered to 16 native Korean speakers living in Germany (age range 27–52; 9 females, 7 males) who were instructed to transcribe the elephant imitations that they were going to hear, using the Korean alphabet, but not being bound by the existence of Korean words. All participants had heard of Koshik, but none of them had experienced him vocalizing or recalled any of his supposed vocabulary. Each auditory stimulus consisted of three identical copies of the same sound, played one after the other, with a total duration of about 4 s per stimulus. To familiarize participants with the sounds, the subjects first heard a random selection of the stimuli. In the test phase, the stimuli were presented in a random order that was identical for all participants. There were 50 stimuli in total, representing 47 unique sounds. The first three were each repeated at a later stage in the experiment (on position 44, 48, and 50, respectively), to serve as controls.

Acoustic Data Analyses

Acoustic analyses were performed using the PRAAT 5.0.29 DSP package. We analyzed 230 speech imitations of Koshik and 100 human model utterances.

The fundamental frequency was measured over the entire utterance with the "to pitch (ac)" command (time step 0.01, time window 0.04 s). The settings for accurate extraction of fundamental frequency were calibrated separately for each sound category (in the case of Koshik's speech

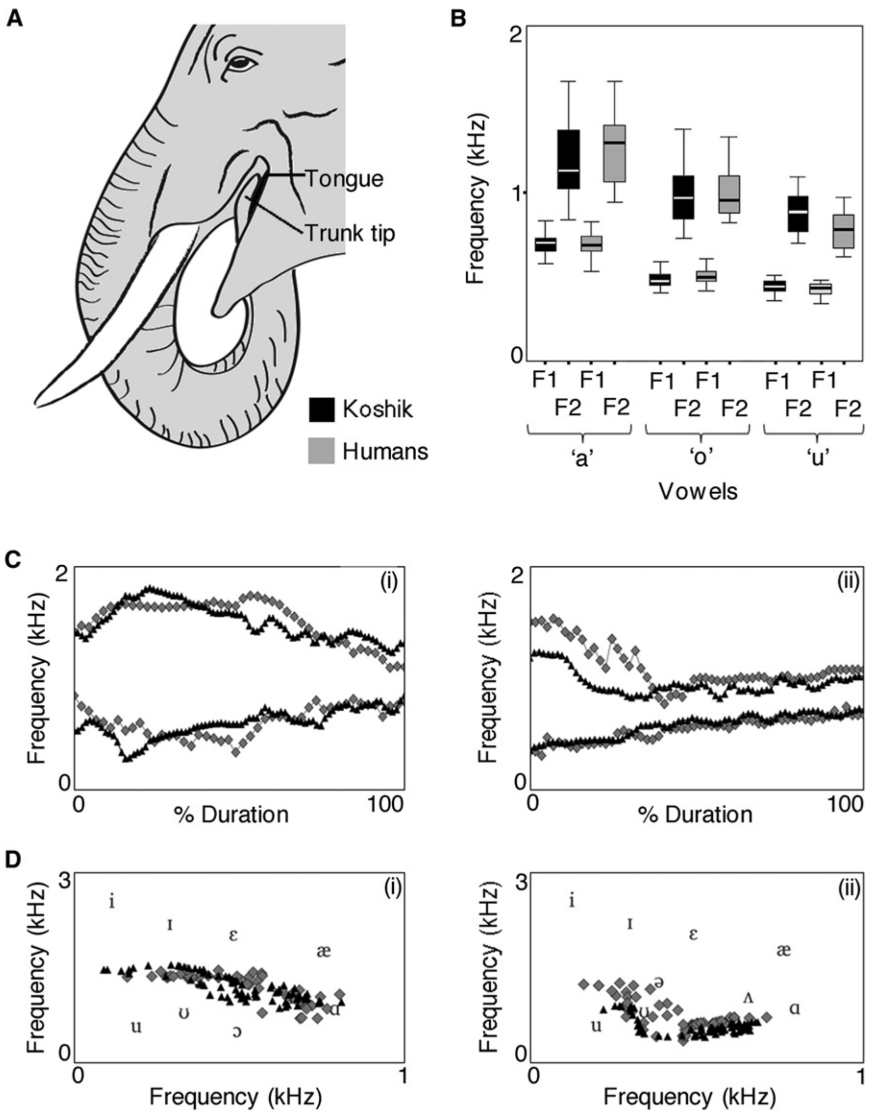

Figure 3. Koshik's Imitation of Human Formant Frequencies

(A) Koshik's posture during speech imitation.

(B) Box plot presentation of the mean peak frequencies of the vowels "a," "o," and "u" of Koshik and his trainers (F1 = Formant 1, F2 = Formant 2).

(C) Time-varying center frequencies of the first two human formants and the corresponding formants of the elephant of (i) "anja" and (ii) "nuo."

(D) Scatterplots of the first formant on the x axis and the second formant on the y axis of the same two utterances as in (C). These data were superimposed upon the mean values for each vowel (given by the phonetic labels) of American English speakers taken from Peterson and Barney [18]. In all cases gray symbols depict human, and black symbols elephant, formant values.

imitations and the human utterances: 120 Hz; rumbles and growls: 10 Hz; squeals, chirps, and trumpets: 300 Hz).

To examine formant frequencies 1 and 2 of the annotated time units of the prospective vowels of Koshik and those of the human models, LPC smoothing was performed on the spectra (number of peaks: 2 in 2000 Hz). The same method was applied to analyze spectral peak frequencies of natural Asian elephant calls (we also used time units of 0.15 s, measured at the middle of the call; if calls were shorter than 0.15 s, we used the entire call; the sampling frequency and the settings were adjusted based on the call type). Due to an inadequate understanding of the production mechanism of particular call types (such as trumpets, squeals, and squeaks/chirps) of the natural Asian elephant vocal repertoire, we use the term "spectral peak" instead of "formant" for those vocalizations.

The timing of the trunk insertion into the oral chamber before vocalization and trunk removal after vocalization was measured from videos using VLC software. The time was noted from the moment that the trunk tip was put into the mouth, until the lower jaw was raised (which was always coincident with the onset of phonation). We further noted the time from the offset of vocalization until the trunk tip was taken out of the trunk. In total, we analyzed a subsample of 50 imitations (including exemplars of all five imitative utterance classes).

Statistical Analyses

All statistical tests were performed in PASW Statistics 18.0. Two-tailed alpha was set to 0.05.

For comparison of acoustic parameters of Koshik's speech imitation with natural Asian elephant calls, we used only tonal Asian elephant calls whose overall spectral structure resembled Koshik's imitation (thus omitting chaotic vocalizations). We analyzed 434 species-typical calls, (rumbles, growls, squeals, chirps, and trumpets).

In order to compare basic acoustic parameters of Koshik's speechlike utterances and the human utterances to natural Asian elephant calls, we entered duration, minimum and maximum fundamental frequency, and the first formant/spectral peak frequency into a discriminant function analysis (DFA). Table S2 gives the number of calls per individual that were analyzed and entered into the DFA. We used a minimum of 5, and a maximum of 10 natural calls per individual. If the data set contained more than 10 calls of an individual, we randomly selected 10 examples out of these data. We tested call membership using the leave-one-out, cross-validation procedure. The results of the DFA are expressed as percentage of classification.

ANOVA and post hoc tests (Bonferroni) were applied to compare the minimum and maximum fundamental frequency of Koshik with the humans and the natural Asian elephant calls.

In addition, ANOVA was also used to compare the means of Koshik's formant frequencies to the corresponding human formants for each vowel. We used the values of the first and the second formant frequency of the most commonly recorded vowels, "a," "o," and "u" (Table S3).

Supplemental Information

Supplemental Information includes three tables, three movies, and four audio files, and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2012.09.022.

Acknowledgments

We are very grateful to the staff of the Everland Zoo for supporting our research. We thank Ji-Hyun Kim-Mietchen for help with analyzing the questionnaire; Jinook Oh for help with the Korean language; Nadja Kavcik for preparing the drawing of Koshik during vocalization; Borami Seo and Kwan-Soo Hong for help in establishing initial contact between the research team and the Everland Zoo; and the Emmen, Walding, Leipzig, and Heidelberg Zoos for providing access to record their elephants. The research was funded by ERC Advanced Grant 230604 "SOMACCA" and a University of Vienna startup grant to W.T.F. and an Austrian Science Fund grant (FWF P23099) to A.S.S.

Received: June 27, 2012
Revised: August 13, 2012
Accepted: September 12, 2012
Published online: November 1, 2012

References

1. Fitch, W.T. (2000). The evolution of speech: a comparative review. Trends Cogn. Sci. *4*, 258–267.
2. Janik, V.M., and Slater, P.J.B. (1997). Vocal learning in mammals. Adv. Stud. Behav. *26*, 59–99.
3. Pepperberg, I.M. (2010). Vocal learning in Grey parrots: A brief review of perception, production, and cross-species comparisons. Brain Lang. *115*, 81–91.
4. Klatt, D.H., and Stefanski, R.A. (1974). How does a mynah bird imitate human speech? J. Acoust. Soc. Am. *55*, 822–832.
5. Marler, P. (1970). A comparative approach to vocal learning: song development in white-crowned sparrows. J. Comp. Physiol. Psychol. *71*, 1–25.
6. Nottebohm, F. (1970). Ontogeny of bird song. Science *167*, 950–956.
7. Brainard, M.S., and Doupe, A.J. (2002). What songbirds teach us about learning. Nature *417*, 351–358.
8. Jarvis, E.D. (2004). Learned birdsong and the neurobiology of human language. Ann. N Y Acad. Sci. *1016*, 749–777.
9. Payne, K., and Payne, R. (1985). Large-scale changes over 19 years in song of humpback whales in Bermuda. Z. Tierpsychol. *68*, 89–114.
10. Ralls, K., Fiorelli, P., and Gish, S. (1985). Vocalizations and vocal mimicry in captive harbor seals: *Phoca vitulina*. Can. J. Zool. *63*, 1050–1056.
11. Eaton, R.L. (1974). A beluga whale imitates human speech. Carnivore. *2*, 22–23.
12. Beeston, R. (1980). Soviet Zoo has talking elephant. The Daily Telegraph, April 9, 1980. p. 16.
13. Hayes, C. (1951). The Ape in Our House (Oxford: Harper).
14. Jarvis, E.D. (2007). Neural systems for vocal learning in birds and humans: a synopsis. J. Ornithol. *148*, 35–44.
15. Jürgens, U. (1992). On the neurobiology of vocal communication. In Nonverbal Vocal Communication: Comparative and Developmental Approaches, H. Papousek, U. Jürgens, and M. Papousek, eds. (Cambridge: Cambridge University Press), pp. 31–42.
16. Lieberman, P. (1984). The Biology and Evolution of Language (Cambridge, MA: Harvard University Press).
17. Poole, J.H., Tyack, P.L., Stoeger-Horwath, A.S., and Watwood, S. (2005). Animal behaviour: elephants are capable of vocal learning. Nature *434*, 455–456.
18. Peterson, G.E., and Barney, H.L. (1952). Control methods used in a study on the vowels. J. Acoust. Soc. Am. *24*, 175–184.
19. Soltis, J. (2010). Vocal communication in African elephants (*Loxodonta africana*). Zoo Biol. *29*, 192–209.
20. Herbst, C.T., Stoeger, A.S., Frey, R., Lohscheller, J., Titze, I.R., Gumpenberger, M., and Fitch, W.T. (2012). How low can you go? Physical production mechanism of elephant infrasonic vocalizations. Science *337*, 595–599.
21. Nair, S., Balakrishnan, R., Seelamantula, C.S., and Sukumar, R. (2009). Vocalizations of wild Asian elephants (*Elephas maximus*): structural classification and social context. J. Acoust. Soc. Am. *126*, 2768–2778.
22. de Silva, S. (2010). Acoustic communication in the Asian elephant, *Elephas maximus maximus*. Behaviour *147*, 825–852.
23. Wemmer, C., and Mishra, H.R. (1982). Observational learning by an Asiatic elephant of an unusual sound production method. Mammalia *46*, 556–557.
24. Hardus, M.E., Lameira, A.R., Van Schaik, C.P., and Wich, S.A. (2009). Tool use in wild orang-utans modifies sound production: A functionally deceptive innovation? Proc. Biol. Sci. *276*, 3689–3694.
25. Patel, A.D., Iversen, J.R., Bregman, M.R., and Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. Curr. Biol. *19*, 827–830.
26. Schachner, A., Brady, T.F., Pepperberg, I.M., and Hauser, M.D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. Curr. Biol. *19*, 831–836.
27. Poole, J.H. (1987). Rutting behaviour in African elephants: the phenomenon of musth. Behaviour *102*, 283–316.
28. Pepperberg, I.M., Naughton, J.R., and Banta, P.A. (1998). Allospecific vocal learning by grey parrots (*Psittacus erithacus*): a failure of videotaped instruction under certain conditions. Behav. Processes *42*, 139–158.
29. Amsler, M. (1947). An almost human grey parrot. Avicult. Mag. *53*, 68–69.
30. West, M.J., Stroud, A.N., and King, A.P. (1983). Mimicry of the human voice by European starlings: the role of social interaction. Wilson Bull. *95*, 635–640.
31. Maglio, V.J. (1973). Origin and evolution of the Elephantidae. Trans. Am. Phil. Soc. *63*, 1–149.
32. Tyack, P.L. (2008). Convergence of calls as animals form social bonds, active compensation for noisy communication channels, and the evolution of vocal learning in mammals. J. Comp. Psychol. *122*, 319–331.
33. Tyack, P.L. (2003). Dolphins communicate about individual-specific social relationships. In Animal Social Complexity: Intelligence, Culture and Individualized Societies, F. de Waal and P.L. Tyack, eds. (Cambridge, MA: Harvard University Press), pp. 342–361.
34. Wittemeyer, G., Douglas-Hamilton, I., and Getz, W.M. (2005). The socioecology of elephants: analysis of the processes creating multitiered social structures. Anim. Behav. *69*, 1357–1371.
35. de Silva, S., Ranjeewa, A.D.G., and Kryazhimskiy, S. (2011). The dynamics of social networks among female Asian elephants. BMC Ecol. *11*, 17.