

Current Biology

Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees

Highlights

- We provide the first evidence for vocal learning of a referential call in non-humans
- Immigrant chimpanzees modify referential food call structure to match hosts' calls
- Call convergence only occurred once affiliative social relationships were formed
- Call structure was not tied to arousal as calls changed while preferences stayed stable

Authors

Stuart K. Watson,
Simon W. Townsend, ..., Victoria West,
Katie E. Slocombe

Correspondence

simon.townsend@ieu.uzh.ch

In Brief

Watson et al. show socially mediated changes in the structure of chimpanzee food calls. This is the first example of vocal learning in referential vocalizations of any non-human species, and it dispels the myth that the structure of such calls is fixed and tied to arousal. This sheds new light on the evolutionary history of human referential words.



Vocal Learning in the Functionally Referential Food Grunts of Chimpanzees

Stuart K. Watson,¹ Simon W. Townsend,^{2,*} Anne M. Schel,¹
Claudia Wilke,¹ Emma K. Wallace,¹ Leveda Cheng,¹
Victoria West,³ and Katie E. Slocombe¹

¹Department of Psychology, University of York, York, North Yorkshire YO10 5DD, UK

²Institute of Evolutionary Biology and Environmental Studies, University of Zurich, 8057 Zurich, Switzerland

³Department of Psychology and Neuroscience, University of St Andrews, St Andrews, Fife KY16 9JP, UK

Summary

One standout feature of human language is our ability to reference external objects and events with socially learned symbols, or words. Exploring the phylogenetic origins of this capacity is therefore key to a comprehensive understanding of the evolution of language. While non-human primates can produce vocalizations that refer to external objects in the environment, it is generally accepted that their acoustic structure is fixed and a product of arousal states [1]. Indeed, it has been argued that the apparent lack of flexible control over the structure of referential vocalizations represents a key discontinuity with language [2]. Here, we demonstrate vocal learning in the acoustic structure of referential food grunts in captive chimpanzees. We found that, following the integration of two groups of adult chimpanzees, the acoustic structure of referential food grunts produced for a specific food converged over 3 years. Acoustic convergence arose independently of preference for the food, and social network analyses indicated this only occurred after strong affiliative relationships were established between the original subgroups. We argue that these data represent the first evidence of non-human animals actively modifying and socially learning the structure of a meaningful referential vocalization from conspecifics. Our findings indicate that primate referential call structure is not simply determined by arousal and that the socially learned nature of referential words in humans likely has ancient evolutionary origins.

Results and Discussion

Language relies on a suite of cognitive capacities that make it far more complex than any other animal communication system. One notable capacity is our ability to label external objects and events with acoustically distinct, referential words. Crucially, the acoustic structure of referential words is socially learned and agreed upon within a community of language users [3–5].

Given the central importance of learning referential words in language, a key question is whether this is a uniquely human trait [6, 7]. Studying referential vocalizations in extant non-human primates (“primates”) with whom we share common ancestors is one approach that can shed light on evolutionary continuity in this capacity [8, 9]. Numerous studies have shown

that primates can refer to external objects in the environment, such as food or predators, with acoustically distinct vocalizations, and receivers respond in ways that suggest the calls are putatively meaningful [10–12]. However, in all cases, the relationship between the acoustic structure of the call and the referent eliciting it appears to be tightly fixed [2]. This is unlikely to be because primates are incapable of flexible modification of vocalizations. Indeed, several primate and non-primate species have been shown to adapt the acoustic structure of social-contact vocalizations in a learned way [9, 13–16], potentially to improve social bonding between individuals [17–23]. Instead, it has been posited that the contexts that elicit referential calls induce arousal states in producers that are associated with specific call structures [2, 24, 25]. One major consequence of this apparent lack of flexibility in the structure of functionally referential calls is the conclusion that this type of call is of little use for understanding the evolutionary origins of language [2].

The aim of our study was to investigate the degree of control and flexibility adult captive chimpanzees have over the structure of their functionally referential food calls [11]. Chimpanzee food grunts are produced almost exclusively in feeding contexts and are directed at valuable social partners [26]. The acoustic structure of these calls varies systematically with the caller’s preference for the referent, which, in captivity, produces food-specific calls [27], and listeners can extract meaningful information about the value and type of food available from group members’ grunts (K.E.S., T. Kaller, J. Call, and K. Zuberbühler, unpublished data; [11]). However, the extent to which vocal learning processes can influence the acoustic structure of chimpanzee referential food calls or exactly how group members converge on a distinct call for a specific food type remains unknown.

In order to address these questions, we investigated the structure of food grunts before and after a rare integration of two groups of captive adult chimpanzees (Table S1) at Edinburgh Zoo in the UK [28]. Acoustic recordings of grunts produced in response to apples were taken before integration in 2010, then after integration in 2011, and again in 2013. We obtained these data from six Edinburgh Zoo residents (ED) and seven immigrant individuals from Beekse Bergen Safari Park (BB) in the Netherlands. We also collected data on the social behavior of all individuals across years to examine whether any acoustic convergence may be related to social cohesion of the integrated group.

We predicted that if acoustic convergence in the food grunts of both subgroups occurred, it would coincide with the development of close social affiliations between subgroups [17, 18]. Systematic preference data for apples were collected in all 3 years in order to explore the possibility that any changes in the acoustic structure of food grunts may simply reflect changes in preference value of the referents and accompanying arousal states [27].

Acoustic Convergence between Two Groups of Chimpanzees

Four acoustic measures extracted from food grunts recorded over the 3 years (2010, 2011, and 2013; see Supplemental

*Correspondence: simon.townsend@ieu.uzh.ch



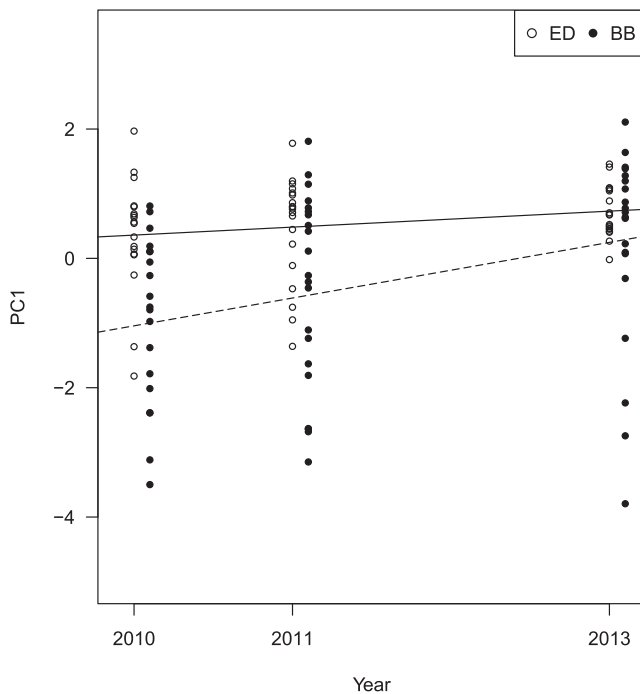


Figure 1. Plot of PC1 against Year

The interaction between year (2010, 2011, 2013) and group (ED, BB) as a function of PC1. Points represent raw data, and lines represent model predictions derived from the GLMMs (dashed lines represent BB; solid lines represent ED).

Information and Table S2) were entered into a principal-component analysis (PCA), where the first principle component (PC1) explained the majority of the total variation: 47%, compared to 26%, 20%, and 7% for components 2–4, respectively. We therefore focused on PC1 for remaining modeling analyses (see Supplemental Information and Table S3 for loading values of the individual acoustic parameters). Generalized linear mixed-effects models (GLMMs) demonstrated that there was a significant interaction between group (BB or ED) and year on PC1 (likelihood ratio [LR] test: $\chi^2 = 4.98$, $df = 1$, $p = 0.025$), suggesting that the effect of year on PC1 differed significantly between the groups. Figure 1 indicates the direction of the interaction with BB individuals (dotted line) converging on PC1 of ED (solid line), which remained stable across years. Assessment of 95% confidence intervals (CIs) indicated that the acoustic structures of food grunts from the two groups were significantly different in 2010 and 2011 (CIs exclude zero; Table S4) but were no longer significantly different in 2013 and thus had converged (CIs include zero; Table S4).

Post hoc GLMMs on the four acoustic variables used to generate the PCA were subsequently employed to investigate the nature of this acoustic change in the BB calls between 2010 and 2013. In 2010, the BB chimpanzees showed a significantly stronger preference for apples compared to the ED chimpanzees (Figure 2), and, in line with previous research [27], the BB calls were higher in frequency than ED calls (Table 1). In 2013, the BB calls had a significantly lower first formant frequency and peak frequency compared to calls in 2010 (mean frequency of the first formant [F1]: LR test, $\chi^2 = 18.1$, $df = 1$, $p < 0.001$; peak frequency: LR test, $\chi^2 = 8.6$, $df = 1$, $p = 0.003$; Table S4) and tended to have a shorter

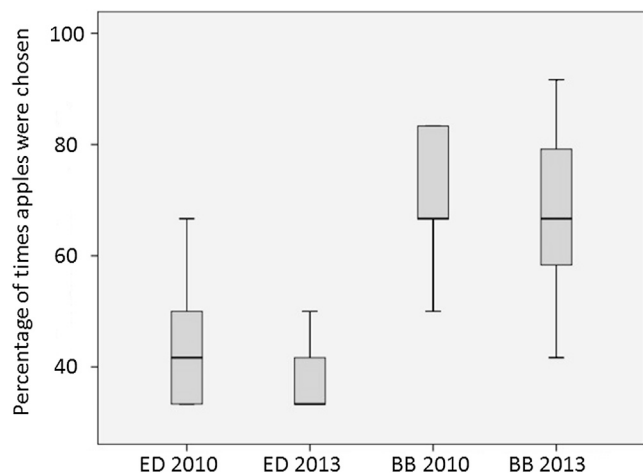


Figure 2. Boxplot of Apple Preferences in 2010 and 2013

Boxplot representing the median percentage of times apples were chosen in pairwise comparisons with five other foods by individuals of ED and BB groups in 2010 and 2013. Data for individuals who also contributed acoustic data from each group (ED = 6; BB = 7) are plotted.

intercall interval (LR test, $\chi^2 = 3.3$, $df = 1$, $p = 0.07$; Table S4). No effect of year was found on the duration of BB calls (LR test, $\chi^2 = 0.02$, $df = 1$, $p = 0.88$; Table S4). This pattern of acoustic change was independent of the BB individuals' preferences for apples, which remained unchanged across years (Figure 2).

Acoustic Convergence Coincides with Social Integration

Using an association index as a measure of social relationships (see Supplemental Information), social network analyses (SNAs) showed that in 2010 and 2011, BB and ED individuals had formed two distinct social subgroups, with maximum modularities of 40% and 30%, respectively [28]. In contrast, in 2013, the SNA showed a maximum modularity of 17% (Figure S1), which is below the 30% threshold for a significant division of data into subgroups [29]. This indicates that, by 2013, the distinct social subgroups identified in 2010 and 2011 by Schel et al. [28] no longer existed, and the group was fully integrated (Figure 3).

Changes in Acoustic Structure Are Not a Result of Changes in Food Preferences

Preference tests demonstrated that the perceived value of apples to ED and BB chimpanzees remained stable across years. BB chimpanzees chose apples over other foods significantly more often than ED chimpanzees in both 2010 (Mann-Whitney U test = 3.50, $z = -2.58$, $p = 0.01$; Figure 2) and 2013 (Mann-Whitney U test = 2.00, $z = -2.76$, $p = 0.004$; Figure 2). Moreover, we found no significant difference in preference value of apples between 2010 and 2013 for either ED (Wilcoxon signed-ranks test $z = -1.633$, $p = 0.25$, $r = -0.66$; Figure 2) or BB (Wilcoxon signed-ranks test $z = -2.10$, $p = 0.875$, $r = -0.79$; Figure 2) individuals.

Conclusions

By comparing the acoustic structure of food grunts given to a specific food type (apples) from two different chimpanzee groups before and after social integration, we demonstrate that these vocalizations are not acoustically fixed and can be modulated independently of the preference value of the

Table 1. Mean Values from the Acoustic Parameters Extracted from the Grunts of Individuals in the ED and BB Groups in 2010 and 2013

	2010	2013
F1		
ED	746	696
BB	900	728
Peak Frequency		
ED	657	597
BB	932	708
Intercall Interval		
ED	0.51	0.40
BB	0.97	0.44
Call Duration		
ED	0.09	0.10
BB	0.13	0.13

referent. Specifically, by 2013, the immigrant BB subgroup had converged on the acoustic structure of food grunts for apples produced by ED individuals (Figures 1 and S2). Furthermore, mere exposure to the new group's different calls for apples was insufficient to trigger convergence: after 1 year of living together, there was no evidence of convergence. It was only in 2013, when social integration was complete and strong social relationships had formed between members of the original subgroups (Figure 3), that BB individuals produced calls similar in structure to those of ED individuals. These results indicate that adult chimpanzees are both able and motivated to modify the acoustic structure of their referential food grunts to match those of unrelated close social partners. Previous observational studies have shown similar learning-based acoustic modulation in the social calls of primates [17–23] and non-primates [14–16]. However, to our knowledge, this is the first example of vocal learning from conspecifics having an influence on the structure of functionally referential vocalizations of any non-human species [2, 9, 30]. This challenges long-held assumptions that, unlike human referential words, functionally referential primate calls cannot be decoupled from the arousal state experienced by the signaler and are completely fixed in their acoustic structure [2, 24].

The acoustic convergence was asymmetrical in its nature, with BB individuals converging on the acoustic structure of calls produced by the resident ED chimpanzees. Several potential factors could account for this effect. First, it is possible that rather than a generalized network of convergence, BB individuals were simply converging on the vocalizations of the dominant male, an ED individual. However, this would be inconsistent with previous acoustic convergence work suggesting that subordinate chimpanzees were equally likely to converge on the pant-hoots of dominant and subordinate males [22]. An alternative is that conformity mechanisms may have motivated the immigrants to adopt the vocal norms of the host group. Although such mechanisms have not yet been confirmed for the learning of primate vocalizations, studies of wild vervet monkeys have shown immigrant males conforming to the normative foraging habits of their host group [31].

Although the chimpanzee subgroups did eventually converge their food grunts, this took approximately 3 years. The slow rate of adoption of the “local” vocal label for apples by the BB individuals can be explained by the long time period

it took for close affiliative relationships to emerge between the subgroups (the integration of two adult, mixed-sex groups would not occur in the wild). In the future, where neighboring wild communities are studied, tracking changes in the acoustic structure of food grunts produced by immigrant females before and after emigration from their natal group would add valuable insights into the time course of vocal learning in a natural setting.

Determining which factors motivate vocal learning of functionally referential calls in non-human primates may also help in understanding what adaptive benefit this ability confers. Given that subtle variation in call structure can transfer different information to listeners [10, 27], it may be that vocal learning facilitates effective communication through ensuring that signalers converge on the same acoustic structure for a particular external event or context. One experimental means of examining this would be to carry out playback experiments before and after an immigrant converges on their host group's calls. If receivers are only capable of decoding the referent of a call after convergence has taken place, this would suggest that the purpose of vocal learning is to facilitate communication. However, if individuals are able to decode the referent of a call prior to acoustic convergence, this might suggest that convergence primarily serves a social function, as has been suggested by previous studies for convergence in non-referential contact calls [22, 23, 32, 33]. Until further experimental evidence is available, it seems most parsimonious to assume social facilitation motivated the observed acoustic changes.

Our study demonstrates that chimpanzees have some control over the acoustic structure of their functionally referential vocalizations and that they are motivated to deploy this ability to make their calls more similar to those of close social partners. Although the adaptive benefits of such behavior need further investigation, the salient finding is that functionally referential food calls are not rigidly coupled to the arousal state of signalers and are open to vocal learning processes. This suggests that such referential calls share more characteristics with referential words in humans than previously thought and indicates the cognitive building blocks underlying the flexible, socially mediated learning of referential words in language may be evolutionarily older than previously thought.

Experimental Procedures

Participants and Study Site

Participants were 18 adult chimpanzees housed at the Budongo Trail in Edinburgh Zoo in the UK. Nine individuals belonged to the “native” ED subgroup and nine to the immigrant BB subgroup, with six and seven of these individuals, respectively, contributing acoustic and preference data (see Supplemental Information). Ethical permission was obtained from the Department of Biology Ethics Committee of the University of York.

Data Collection

Acoustic recordings of food grunts were obtained during group apple feeds several times per week during three data collection periods: May 2010–August 2010, April 2011–July 2011, and April 2013–July 2013. Only good-quality recordings in which the identity of the caller was unambiguous were analyzed.

To ensure direct comparability to the social data reported by Schel et al. [28] in 2010 and 2011, identical measures were taken. Thus, social data constituted instantaneous scan samples of group composition and point samples of the identity of the focal individual's nearest neighbor, which were collapsed into a single composite association measure on which SNAs were performed (see Supplemental Information).

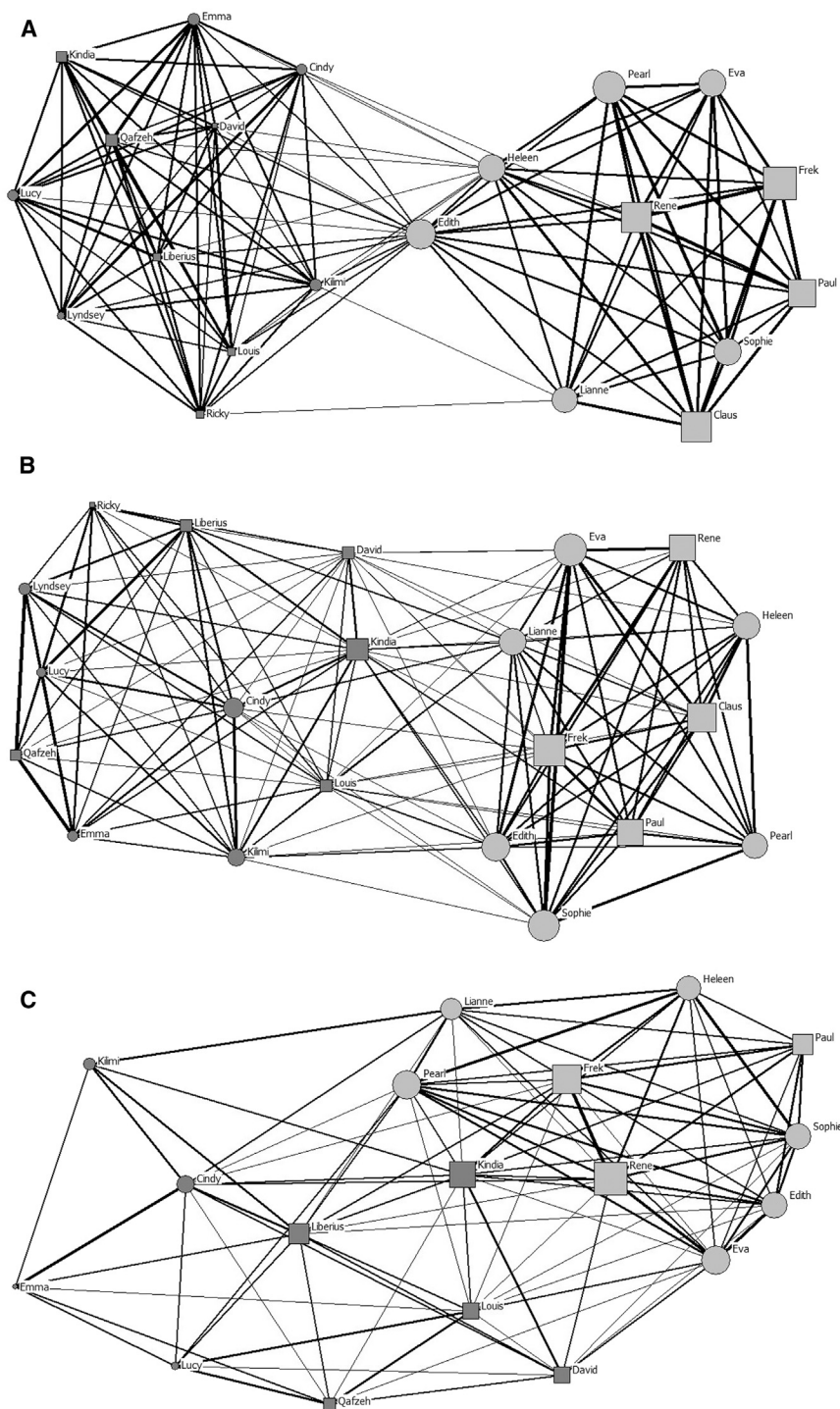


Figure 3. Sociograms from 2010, 2011, and 2013 (A–C) Sociograms illustrating association patterns between July 2010–December 2010 (A), April 2011–October 2011 (B) (data from Schel et al. [28]), and April 2013–July 2013 (C). ED individuals are in dark gray; BB individuals are in light gray. Males are shown as squares, and females are shown as circles. The size of the node scales according to eigenvector centrality index for that individual. Thickness of lines scale with the strength of association. All relationships of less than the community mean ($A + B = 0.05$, $C = 0.15$) were removed for clarity.

Food preference data were collected by presenting chimpanzees with pairwise choices between apples and five other types of food (see [Supplemental Information](#)). Each individual's preference value for apples was determined as the percentage of food types that they rejected in favor of apples.

Statistical Analysis

GLMMs were used to investigate the influence of explanatory variables (e.g., year and social group) on the acoustic structure of food grunts. Because we had repeated measures (call bouts) from each individual, we fitted subject as a random factor using the package lme4. All analyses were conducted in R v.0.98.1062.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, two figures, and four tables and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2014.12.032>.

Acknowledgments

We are grateful to the Royal Zoological Society of Scotland for permission to conduct research at Edinburgh Zoo and to the Budongo Trail keepers for their help. We thank Andri Manser for statistical advice and analysis scripts; Rachel Harrison for contributing to social data; and three anonymous

reviewers for their constructive comments on earlier versions of the manuscript. This research was funded by a BBSRC grant to K.E.S., and S.W.T. was funded by the University of Zurich.

Received: October 1, 2014

Revised: November 20, 2014

Accepted: December 10, 2014

Published: February 5, 2015

References

- Tomasello, M. (2008). *Origins of Human Communication* (Cambridge: MIT Press).
- Wheeler, B.C., and Fischer, J. (2012). Functionally referential signals: a promising paradigm whose time has passed. *Evol. Anthropol.* *21*, 195–205.
- Baldwin, D.A., Markman, E.M., Bill, B., Desjardins, R.N., Irwin, J.M., and Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Dev.* *67*, 3135–3153.
- Kirby, S., and Hurford, J.R. (2002). The emergence of linguistic structure: an overview of the iterated learning model. In *Simulating the Evolution of Language*, A. Cangelosi and D. Parisi, eds. (London: Springer), pp. 121–147.
- Smith, K., Kirby, S., and Brighton, H. (2003). Iterated learning: a framework for the emergence of language. *Artif. Life* *9*, 371–386.
- Hurford, J.R. (2007). *The Origins of Meaning: Language in the Light of Evolution* (Oxford: Oxford University Press).
- Fitch, W.T., Huber, L., and Bugnyar, T. (2010). Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* *65*, 795–814.
- Hauser, M.D., Chomsky, N., and Fitch, W.T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* *298*, 1569–1579.
- Fedurek, P., and Slocombe, K.E. (2011). Primate vocal communication: a useful tool for understanding human speech and language evolution? *Hum. Biol.* *83*, 153–173.
- Seyfarth, R.M., Cheney, D.L., and Marler, P. (1980). Vervet monkey alarm calls: semantic communication in a free-ranging primate. *Anim. Behav.* *28*, 1070–1094.
- Slocombe, K.E., and Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Curr. Biol.* *15*, 1779–1784.
- Zuberbühler, K., Noë, R., and Seyfarth, R.M. (1997). Diana monkey long-distance calls: messages for conspecifics and predators. *Anim. Behav.* *53*, 589–604.
- Janik, V., and Slater, P.B. (1997). Vocal learning in mammals. *Adv. Stud. Behav.* *26*, 59–99.
- Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence. *Anim. Behav.* *37*, 64–73.
- Hile, A.G., Plummer, T.K., and Striedter, G.F. (2000). Male vocal imitation produces call convergence during pair bonding in budgerigars, *Melopsittacus undulatus*. *Anim. Behav.* *59*, 1209–1218.
- Brown, E.D., and Farabaugh, S.M. (1997). What birds with complex social relationships can tell us about vocal learning: vocal sharing in avian groups. In *Social Influences on Vocal Development*, C.T. Snowdon and M. Hausberger, eds. (Cambridge: Cambridge University Press), pp. 98–127.
- Geissmann, T. (1986). Mate change enhances duetting activity in the siamang gibbon (*Hylobates syndactylus*). *Behaviour* *96*, 17–27.
- Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biol. Rev. Camb. Philos. Soc.* *77*, 57–76.
- Elowson, A.M., and Snowdon, C.T. (1994). Pygmy marmosets, *Cebuella pygmaea*, modify vocal structure in response to changed social environment. *Anim. Behav.* *47*, 1267–1277.
- Snowdon, C.T., and Elowson, A.M. (1999). Pygmy marmosets modify call structure when paired. *Ethology* *105*, 893–908.
- Marshall, A.J., Wrangham, R.W., and Arcadi, A.C. (1999). Does learning affect the structure of vocalizations in chimpanzees? *Anim. Behav.* *58*, 825–830.
- Mitani, J.C., and Brandt, K.L. (1994). Social factors influence the acoustic variability in the long-distance calls of male chimpanzees. *Ethology* *96*, 233–252.
- Mitani, J.C., and Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: tests of three hypotheses. *Behaviour* *135*, 1041–1064.
- Morton, E.S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *Am. Nat.* *111*, 855–869.
- Rendall, D., Owren, M.J., and Ryan, M.J. (2009). What do animal signals mean? *Anim. Behav.* *78*, 233–240.
- Schel, A.M., Machanda, Z., Townsend, S.W., Zuberbühler, K., and Slocombe, K.E. (2013). Chimpanzee food calls are directed at specific individuals. *Anim. Behav.* *86*, 955–965.
- Slocombe, K.E., and Zuberbühler, K. (2006). Food-associated calls in chimpanzees: responses to food types or food preferences? *Anim. Behav.* *72*, 989–999.
- Schel, A.M., Rawlings, B., Claidière, N., Wilke, C., Wathan, J., Richardson, J., Pearson, S., Herrelko, E.S., Whiten, A., and Slocombe, K. (2013). Network analysis of social changes in a captive chimpanzee community following the successful integration of two adult groups. *Am. J. Primatol.* *75*, 254–266.
- Newman, M.E. (2004). Analysis of weighted networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* *70*, 056131.
- Hollen, L.I., and Radford, A.N. (2009). The development of alarm call behaviour in mammals and birds. *Anim. Behav.* *78*, 791–800.
- van de Waal, E., Borgeaud, C., and Whiten, A. (2013). Potent social learning and conformity shape a wild primate's foraging decisions. *Science* *340*, 483–485.
- 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.
- Hausberger, M., Richard-Yris, M.-A., Henry, L., Lepage, L., and Schmidt, I. (1995). Song sharing reflects the social organization in a captive group of European starlings (*Sturnus vulgaris*). *J. Comp. Psychol.* *109*, 222–241.