Published in final edited form as:

Am J Primatol. 2015 November ; 77(11): 1143–1148. doi:10.1002/ajp.22449.

Multimodal Communication in Chimpanzees

JARED P. TAGLIALATELA^{1,2,*}, JAMIE L. RUSSELL^{2,3}, SARAH M. POPE³, TAMARA MORTON¹, STEPHANIE BOGART³, LISA A. REAMER⁴, STEVEN J. SCHAPIRO⁴, and WILLIAM D. HOPKINS^{2,3,4}

¹Department of Ecology, Evolution, and Organismal Biology, Kennesaw State University, Kennesaw, Georgia

²Division of Developmental and Cognitive Neuroscience, Yerkes National Primate Research Center, Atlanta, Georgia

³Neuroscience Institute and Language Research Center, Georgia State University, Atlanta, Georgia

⁴Department of Veterinary Sciences, Michale E. Keeling Center for Comparative Medicine and Research, The University of Texas MD Anderson Cancer Center, Bastrop, Texas

Abstract

A fundamental characteristic of human language is multimodality. In other words, humans use multiple signaling channels concurrently when communicating with one another. For example, people frequently produce manual gestures while speaking, and the words a person perceives are impacted by visual information. For this study, we hypothesized that similar to the way that humans regularly couple their spoken utterances with gestures and facial expressions, chimpanzees regularly produce vocalizations in conjunction with other communicative signals. To test this hypothesis, data were collected from 101 captive chimpanzees living in mixed-sex social groupings of seven to twelve individuals. A total of 2,869 vocal events were collected. The data indicate that approximately 50% of the vocal events were produced in conjunction with another communicative modality. In addition, approximately 68% were directed to a specific individual, and these directed vocalizations were more likely to include a signal from another communicative modality than were vocalizations that were not directed to a specific individual. These results suggest that, like humans, chimpanzees often pair their vocalizations with signals from other communicative modalities. In addition, chimpanzees appear to use their communicative signals strategically to meet specific socio-communicative ends, providing support for the growing literature that indicates that at least some chimpanzee vocal signaling is intentional.

Keywords

chimpanzees; vocal communication; multimodal communication; language origins

*Correspondence to: Jared P. Taglialatela, PhD, Assistant Professor of Biology, Department of Biology and Physics, Kennesaw State University, 1000 Chastain Road, Kennesaw, GA 30144 USA. jtaglial@kennesaw.edu.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

INTRODUCTION

Human language is unique within the animal kingdom. However, the study of communicative behavior in extant nonhuman primates is critical for understanding the evolutionary origins of this uniqueness. Specifically, to determine why and how such a sophisticated communication system has evolved in humans, one must distinguish those characteristics that are derived in the human lineage and those that are ancestral. However, human language does not leave direct indelible marks in the fossil record. Therefore, to decipher the evolutionary origins of human language, one must identify similarities, as well as differences, between the communicative behavior of humans and their closest phylogenetic relatives, chimpanzees.

Despite having the ability to rely exclusively on speech sounds, allowing complex communication with individuals that are not in direct view or physical proximity, humans often use multiple signaling modalities concomitantly when communicating with one another. For example, people frequently produce manual gestures while speaking [Kimura, 1973; McNeil, 1992]. In addition, the appearance of a speaker's lips and mouth articulating a certain phoneme influences the speech sound that is perceived by a listener [McGurk and Macdonald, 1976]. Therefore, in terms of both production and perception, human language utilizes more than just the auditory modality. Consistent with this, Broca's area (a region of the human brain that is critical for language) is involved in both speech and American Sign Language production, suggesting that the functional role of this region in mediating language is modality independent [Emmorey et al., 2007]. Therefore, although humans are able to communicate effectively with social partners solely through the vocal channel, signals from other modalities provide important (and often critical) information that may modulate intensity, increase precision, or otherwise alter the message contained within the vocal signal [Partan and Marler, 1999, 2005].

Indeed, a number of recent studies have reported that nonhuman primates, in the wild as well as in captivity, similarly produce communicative signals from multiple modalities simultaneously [Leavens et al., 2005, 2010; Micheletta et al., 2013; Tagliatela et al., 2011], and have noted the importance of adopting an integrated and multimodal approach to the study of primate communication [Micheletta et al., 2013; Slocombe et al., 2011; Tagliatela et al., 2011]. In other words, for those interested in elucidating the phylogenetic origins of human language—a multimodal communicative system—comparative studies involving extant nonhuman primates should similarly employ methods that consider the concurrent production and usage of communicative signals from multiple modalities.

For this study, we hypothesized that similar to the way that humans regularly couple their spoken utterances with signals from other communicative modalities, chimpanzees regularly produce vocalizations in conjunction with other communicative signals. If chimpanzees indeed pair their vocal utterances with signals from other communicative modalities, it would support the hypothesis that human spoken language evolved from a multimodal communication system present in the common ancestor of humans and chimpanzees approximately 6 million years ago. In addition we sought to determine how these other modalities were being used, and specifically, if signals from other modalities were being

used to modulate the vocal signal for particular communicative ends. To this end we predicted that chimpanzees would be more likely to pair their vocal utterances with signals from other modalities when directing these signals to communicative partners in close physical proximity, than when producing vocalizations that are not directed.

MATERIALS AND METHODS

All of the data collected for this study were approved by the appropriate Institutional Animal Care and Use Committees and adhered to the legal requirements of the United States of America. In addition, this research adhered to the American Society of Primatologists principles for the ethical treatment of primates.

To test the hypothesis that chimpanzees would be more likely to pair their vocal utterances with signals from other modalities when directing these signals to communicative partners in close physical proximity, than when producing vocalizations that are not directed, data were collected from 101 chimpanzees (37 males and 64 females, mean age = 25.39 years) housed at the Michale E. Keeling Center for Comparative Medicine and Research (KCCMR) in Bastrop, Texas ($N = 71$; 31 males) and the Yerkes National Primate Research Center Field Station (YNPRC) in Lawrenceville, Georgia ($N = 30$; 6 males). Subjects live in open-top, outdoor corrals with adjoining indoor runs. All subjects were housed in mixed-sex social groupings of seven to twelve individuals. All data were collected from subjects while they were in their outdoor corrals (YNPRC = 2 social groups, outdoor corral A is approximately 7,488 square feet; outdoor corral B is approximately 5,600 square feet; KCCMR = 8 social groups, outdoor corrals are approximately 4,346 square feet).

All data were collected in real-time using an all-occurrence sampling method [Altmann, 1974]. Observations were taken between the hours of 7:30 a.m. and 4:30 p.m. Monday through Friday. For the subjects housed at the KCCMR, data were collected daily, Monday through Friday, between June 1 and June 29, 2011 and represent approximately 110 hrs of observations. For the subjects housed at the YNPRC, data were collected approximately two days per week from May 3, 2011 to December 6, 2011 and from May 2, 2012 to July 26, 2012 and represent approximately 300 hrs of observations.

A vocal event was defined as any vocalization made by an individual chimpanzee that was not directed to a human. In order to be coded, the observer needed to have a clear view of the individual vocalizing and of any other chimpanzees in the enclosure. In some instances, (e.g., multiple chimpanzees were vocalizing at once), it was not possible to determine all of the details of a vocal event. Those instances were not coded. For each clearly visible/audible vocal event, we recorded the date, time, vocalization type (see Table I), whether or not there was a concomitant communicative signal (CCS) and whether or not the vocalization was directed toward another conspecific. Vocalizations were classified into 6 categories (see Table I). A communicative signal was considered concomitant if it happened within the same communicative bout (immediately following or happening at the same time as the vocalization; i.e., approximately 2 sec before or after the vocal event). If a CCS occurred, the real-time coder made a written note describing the observed communicative signal. A second experimenter later compiled these CCS behaviors into one of the 30 types described

in Table II. A vocalization was considered “directed” to a conspecific if the producer emitted the vocalization while either 1) performing a focused look or glance towards another chimpanzee, or 2) approaching another individual. For directed vocal events, the experimenter noted the name of the individual(s) the vocalization was directed towards.

Multiple observers collected data for this study (JPT, JLR, SMP, TM, SB, and LAR). To ensure common classification of vocalizations, CCS, and directedness, observers independently reviewed both real-time and pre-recorded vocal events following a training period before data collection began. In addition, all of the observers, (excluding LAR) collected data from both populations of chimpanzees.

RESULTS

Given that the vocal events data were collected using a continuous, all-occurrence sampling method, the data were initially pooled across subjects. A total of 2,869 vocal events were collected, (mean = 28 vocal events per subject; Supplementary Table I) and the data indicate that 50% ($N = 1,424$) of the vocal events were produced in conjunction with another communicative modality. The proportion of vocal events that included a CCS differed across vocalization types (Fig. 1a; $\chi^2(5) = 679.19$, $P < .001$). In addition, 68% ($N = 1,957$) of the vocal events were directed to another individual. Similar to CCS use, the proportion of vocal events that were directed to a specific individual differed across vocalization types (Fig. 1b; $\chi^2(5) = 1119.58$, $P < .001$).

We were next interested in determining if the use of CCS was related to whether or not a vocalization was directed to a specific conspecific. To evaluate this, we restricted our analyses to those chimpanzees that had a minimum of 10 vocal events ($N = 68$), and performed a within-subject analysis. We found a significant interaction between whether or not a vocalization was directed to a specific individual and if that sound was produced in conjunction with a CCS ($F(1,67) = 17.07$; $P < .001$). Directed vocalizations were more likely to include a CCS (Fig. 2).

DISCUSSION

These data indicate that chimpanzees frequently produce vocalizations in conjunction with signals from other communicative modalities and that a significant majority are directed to a specific individual. Perhaps more importantly, the results suggest that certain types of vocalizations are more likely than others are to include a CCS (Fig. 1a) and to be directed (Fig. 1b). Of particular note is the fact that pants and pant grunts were more frequently produced with a CCS than any other vocal type (~74%). This category includes calls that are particularly important for regulating intra-party relations [Goodall, 1986] and are almost exclusively directed to a specific individual (Fig. 1b). Previously, it has been shown that two broad categories of vocalizations (Proximal and Broadcast vocalizations) are processed differently in the chimpanzee brain [Tagliatalata et al., 2009]. “Proximal” vocalizations (PRV) are relatively low intensity vocalizations typically produced by individuals in close spatial proximity of conspecifics, and are directed toward these individuals. “Broadcast” vocalizations (BCV) are much higher amplitude calls as compared with PRV, and are

produced by individuals in the presence of conspecifics, but appear not to be directed to these individuals. Taglialatela et al. [2009] demonstrated that the perception of PRV, which would include call types such as pants and pant grunts, differentially activated the right posterior temporal lobe, including the planum temporale, when compared to the perception of BCV. These results are consistent with existing literature on nonhuman primate vocal perception, and suggest that vocalizations that are directed to a specific individual in close proximity may be both produced and processed differently than those call types that are used in more “broadcast” type contexts [Ghazanfar et al., 2005; Taglialatela et al., 2009].

Related to this idea, we predicted that multimodal signals (i.e., vocalizations produced in conjunction with signals from other communicative modalities) would occur at higher frequencies when callers were directing their calls to specific individuals. In these instances, the social partner is likely to be looking at the caller and presumably able to benefit from additional visual or tactile signals that accompany the auditory signal. Indeed Ghazanfar et al., (2005) have proposed this directed/non-directed distinction may explain different patterns of responses in the auditory cortex of macaque monkeys following the presentation of two different types of proximal vocalizations (coos and grunts). Our results indicating that directed vocalizations were more likely to include a CCS than those that were not directed to a specific individual are consistent with this prediction. Further, these data suggest that multimodal signals are used strategically to meet specific—and immediate—socio-communicative ends, providing support for the growing literature that indicates that some chimpanzee vocal signaling is intentional [Gruber and Zuberbuhler, 2013; Laporte and Zuberbuhler, 2010; Schel et al., 2013].

Acknowledgments

Contract grant sponsor: National Institute on Deafness and Other Communication Disorders; contract grant number: R15DC011005; contract grant sponsor: Kennesaw State University College of Science and Mathematics; contract grant sponsor: Dr. Nadia Girardot Ethology Scholarship; contract grant sponsor: NIH/NCRR; contract grant number: U42-RR15090.

The authors thank the veterinary and animal care staffs at the Michale E. Keeling Center for Comparative Medicine and Research as well as the Yerkes National Primate Research Center. This research was supported in part by Award no. R15DC011005 from the National Institute on Deafness and Other Communication Disorders to J.P.T., as well as by the Kennesaw State University College of Science and Mathematics and the Dr. Nadia Girardot Ethology Scholarship to S.M.P. Chimpanzee maintenance at the Keeling Center is funded by NIH/NCRR U42-RR15090.

References

- Altmann J. Observational study of behavior: sampling methods. *Behaviour*. 1974; 49:227–267. [PubMed: 4597405]
- Emmorey K, Mehta S, Grabowski TJ. The neural correlates of sign versus word production. *NeuroImage*. 2007; 36:202–208. [PubMed: 17407824]
- Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. *J Neurosci*. 2005; 25:5004–5012. [PubMed: 15901781]
- Goodall, J. *The chimpanzees of Gombe: Patterns of behavior*. Cambridge: Harvard University Press; 1986.
- Gruber T, Zuberbuhler K. Vocal recruitment for joint travel in wild chimpanzees. *PLoS ONE*. 2013; 8:e76073. [PubMed: 24086688]

- Kimura D. Manual activity during speaking—I. Right handers. *Neuropsychologia*. 1973; 11:45–50. [PubMed: 4694780]
- Laporte MNC, Zuberbühler K. Vocal greeting behaviour in wild chimpanzee females. *Anim Behav*. 2010; 80:467–473.
- Leavens DA, Russell JL, Hopkins WD. Intentionality as measured in the persistence and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child Dev*. 2005; 76:291–306. [PubMed: 15693773]
- Leavens DA, Russell JL, Hopkins WD. Multi-modal communication and its social contextual use in captive chimpanzees (*Pan troglodytes*). *Anim Cogn*. 2010; 13:33–40. [PubMed: 19504272]
- McGurk H, Macdonald J. Hearing lips seeing voices. *Nature*. 1976; 264:746–748. [PubMed: 1012311]
- McNeil, D. Hand and mind: What gestures reveal about thought. Chicago: University of Chicago Press; 1992.
- Micheletta J, Engelhardt A, Matthews LEE, et al. Multicomponent and Multimodal Lipsmacking in Crested Macaques (*Macaca nigra*). *Am J Primatol*. 2013; 75:763–773. [PubMed: 23225489]
- Partan S, Marler P. Communication goes multimodal. *Science*. 1999; 283:1272–1273. [PubMed: 10084931]
- Partan SR, Marler P. Issues in the classification of multimodal communication. *Am Nat*. 2005; 166:231–245. [PubMed: 16032576]
- Schel AM, Townsend SW, Machanda Z, et al. Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE*. 2013; 8:e76674. [PubMed: 24146908]
- Slocombe KE, Waller BM, Liebal K. The language void: the need for multimodality in primate communication research. *Anim Behav*. 2011; 81:919–924.
- Tagliatela JP, Russell JL, Schaeffer JA, Hopkins WD. Visualizing vocal perception in the chimpanzee brain. *Cerebral Cortex*. 2009; 19:1151–1157. [PubMed: 18787228]
- Tagliatela JP, Russell JL, Schaeffer JA, Hopkins WD. Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS One*. 2011; 6:e18852. [PubMed: 21533079]

Fig. 1.

(a) Chimpanzee vocalizations ($N = 2869$) produced in conjunction with a CCS ($n = 1424$) by vocalization type. Please see Table I for vocalization types and Table II for descriptions of CCS. (b) Chimpanzee vocalizations ($N = 2869$) directed to a conspecific ($n = 1957$) by vocalization type. Please see Table I for description of vocalization types. AL = Alarm calls; FC = Food Calls; PG = Pants and pant grunts; PH = Pant hoots; SC = Screams and whimpers; OT = Others.

Fig. 2.

Mean (SE) number of vocal events that are produced with and without a CCS that were directed to a particular individual. Chimpanzee vocalizations that were directed to a specific individual were more likely to be produced in conjunction with a CCS ($F(1,67) = 17.07$; $P < .001$). Analysis restricted to subjects with at least 10 vocal events ($N = 68$).

TABLE I

Chimpanzee Vocalization Types

Vocalization type	Description
Pant hoots (PH)	Voiced on both inhalation and exhalation and incorporate a series of “hoo” sounds which may or may not escalate to a climactic scream or piercing “ahh” vocalization.
Alarm calls (AL)	Loud, sharp vocalizations that may sound like “wraa,” or “waa”. Alarms are given in the context of a real or perceived danger (snake, truck, etc.) when a chimpanzee is on high alert. Some chimpanzees may also make a quieter “hoo” sound in this context. Alarms are also often made by bystanders during a fight.
Food calls (FC)	Consist of food barks and grunts and are short, “aaa” vocalizations that are produced by short exhalations. They are often produced in a series and can range in pitch from low to high. Grunts are relatively low frequency and noisy, whereas barks tend to be tonal and higher in frequency than grunts. These calls are typically associated with the anticipation of eating or receiving food or other positive experiences
Pants and pant grunts (PG)	Fast, repetitive low frequency vocalizations made on both inhalation and exhalation. Pants are very quiet and breathy and are sometimes accompanied by placing an open mouth on another individual while panting. Pants can be difficult to hear from a distance and are most easily recognized by the quick, rhythmic movements of the body. Pant grunts are quiet vocalizations but louder than pants and sound like a series of “ohoh” or “uhuh” sounds made in quick succession.
Screams and whimpers (SC)	Screams are loud, high-pitched, voiced shrieks and at their most intense can be raspy or even hoarse sounding. Screams are associated with fear, submission, distress or agitation. Whimpering sounds a bit like modulated, high-pitched “hoo” sounds or crying and often progresses into screams. Whimpering occurs in chimpanzees of all ages during distress or fear and by infants when being weaned.
Others (OT)	Any vocalization that does not fall into one of the above categories was classified as “other.” A description of the sound and context used was included in the notes section of the data sheet.

TABLE II

Concomitant Communicative Signals (CCS)

CCS	Description
Aggressive hit (AH)	Individual uses hand to strike or push recipient with force, pilo-erect.
Bounce (BO)	Subject is in a seated or quadrupedal position and repeatedly bounces body up and down, often with knuckles on the ground with arms stiff.
Cage bang (CB)	Subject uses hand to forcefully strike a substrate with the apparent intention of creating noise.
Chase (CH)	Subject quickly and purposefully pursues the recipient of the vocalization.
Clap (CL)	Subject contacts hands together forcefully to create noise.
Crouch (CR)	Subject places body in a submissive position with chest and stomach moving down towards the ground with elbows fully bent
Display (DI)	Subject is pilo-erect while swaying, charging, slapping and/or pushing objects in enclosure.
Embrace (EM)	Individual places one or two arms around another, generally around his or her back from the front but can be from the back around his or her middle.
Fear Grimace (FG)	Subject's mouth is widely or partially opened with the corners fully withdrawn and lips retracted from upper and lower teeth and most of the gums
Facial Inspect (FI)	Subject brings his/her face within 4 inches of the recipient's face using a directed gaze that lasts at least 3 sec
Manual gesture (GE)	Subject produces a manual signal that does not fall into any of the other gesture categories.
Genital Inspect (GI)	Subject touches recipient's swelling or penis with fingertip(s) or hand.
Grab (GR)	Subject uses his/her hand(s) to forcefully grasp recipient.
Head bob (HB)	Subject moves head in an up and down motion repeatedly.
Kiss (KI)	Subject places a closed mouth with pursed lips on another's body part.
Mount (MO)	Subject contacts another individual from behind with their midsection and moves pelvis repetitively, can be between individuals of the same or opposite sex, may or may not include sexual contact.
Open mouth kiss (OM)	Subject gently places open mouth around a body part of another individual
Play (PL)	Subject is engaged in tickling, chasing, wrestling, etc. with another individual, often accompanied by laughing and/or play face.
Present (PR)	Subject turns to present swelling or rump to another individual.
Rocking (RO)	Subject shifts own weight repeatedly from side to side, usually in a seated position.
Self hit (SH)	Subject crosses one or both hands across torso and slaps their own body repeatedly.
Swagger/sway (SW)	Subject is pilo-erect and moves torso from side to side, often bipedal.
Threat gesture (TG)	Includes arm raises (subject swings arm in a quick, upward motion toward another individual with palm facing down), flaps (subject raises one arm and hand and makes a downward slapping movement of the hand, typically in the direction of another individual—no forceful contact with substrate) and wrist shakes (subject shakes the hand vigorously and repeatedly with a flexible wrist toward another individual with arm mostly to fully extended).
Throw (TH)	Subject uses hand to toss debris (dirt, feces, etc.).
Touch (TO)	Subject makes any sort of contact with another individual with the front or back of their hand or fingers, without appreciable force, that does not fall into one of the more specific categories of touch.
Mouth on wrist/finger (WM)	Subject places his/her mouth around the recipient's wrist or finger.
Wrist/finger present (WP)	Subject flexes the wrist while holding the back or side of hand out toward another individual, may include placing a finger or hand into another individual's mouth. Note whether or not contact occurs.
AUD	Any auditory communicative signal that does not fit into one of the more specific categories listed above.
TAC	Any tactile communicative signal that does not fit into one of the more specific categories listed above.
VIS	Any visual communicative signal that does not fit into one of the more specific categories listed above.