

Functional flexibility of infant vocalization and the emergence of language

D. Kimbrough Oller^{a,b,c,1}, Eugene H. Buder^{a,b}, Heather L. Ramsdell^d, Anne S. Warlaumont^e, Lesya Chorna^f, and Roger Bakeman^g

^aSchool of Communication Sciences and Disorders, University of Memphis, Memphis, TN 38105; ^bInstitute for Intelligent Systems, University of Memphis, Memphis, TN 38152; ^cThe Konrad Lorenz Institute for Evolution and Cognition Research, A-3422 Altenberg, Austria; ^dDepartment of Communication Sciences and Disorders, Idaho State University, Pocatello, ID 83209; ^eCognitive and Information Sciences, University of California, Merced, CA 95343; ^fMathematics Department, Milwaukee School of Engineering, Milwaukee, WI 53202; and ^gDepartment of Psychology, Georgia State University, Atlanta, GA 30303

Edited* by E. Anne Cutler, Max Planck Institute for Psycholinguistics, Heilig Landstichting, The Netherlands, and approved March 7, 2013 (received for review January 8, 2013)

We report on the emergence of functional flexibility in vocalizations of human infants. This vastly underappreciated capability becomes apparent when prelinguistic vocalizations express a full range of emotional content—positive, neutral, and negative. The data show that at least three types of infant vocalizations (squeals, vowel-like sounds, and growls) occur with this full range of expression by 3–4 mo of age. In contrast, infant cry and laughter, which are species-specific signals apparently homologous to vocal calls in other primates, show functional stability, with cry overwhelmingly expressing negative and laughter positive emotional states. Functional flexibility is a sine qua non in spoken language, because all words or sentences can be produced as expressions of varying emotional states and because learning conventional “meanings” requires the ability to produce sounds that are free of any predetermined function. Functional flexibility is a defining characteristic of language, and empirically it appears before syntax, word learning, and even earlier-developing features presumed to be critical to language (e.g., joint attention, syllable imitation, and canonical babbling). The appearance of functional flexibility early in the first year of human life is a critical step in the development of vocal language and may have been a critical step in the evolution of human language, preceding protosyntax and even primitive single words. Such flexible affect expression of vocalizations has not yet been reported for any non-human primate but if found to occur would suggest deep roots for functional flexibility of vocalization in our primate heritage.

evolution of language | infant communication | flexibility in communication | language development | primate communication

Research on evolution and development of language has been devoted primarily to syntax, the uniquely human capacity to produce well-formed complex sentences (1–4). Additional work has targeted the emergence of simpler communicative structures and thus has shifted attention back in evolutionary time to an earlier possible split of hominins from the primate background. For example, research has considered the presumably earlier evolution of simple sentences or “protosyntax” (5). Other work influenced by recent trends in evolutionary developmental biology (evo-devo) (6, 7) has focused on infrastructure for language, invoking capabilities logically more foundational even than protosyntax and presumably moving the communicative differentiation of hominins from other primates much farther back. For example, symbolic expression in single words beginning in modern human development at about 12 mo is a precursor to even the simplest syntax (8). Moving the evolutionary focus even farther back in time, joint attention—infant pointing with gaze alternation between an object and an adult interactor, occurring before the end of the first year—is deemed a critical precursor to words (9). Similarly, canonical babbling (onset at about 7 mo) is a crucial step toward verbal vocabulary because development and imitation of canonical syllables (e.g., “baba”) is required for extensive word learning (10, 11). Recent reasoning influenced by evo-devo converges on the contention that the evolution of language in hominins and the development of language in modern infants are

guided by a common set of infrastructural principles, with similar foundational steps (such as canonical babbling, joint attention, and so on) being required in both cases for subsequent language-related advancement (12–14).

We continue to be influenced by evo-devo logic and seek the deepest foundations in our communicative phylogenetic history and the earliest points of departure in evolution and development between humans and other animals, especially our closest primate relatives. Our report focuses on a capability appearing earlier in human development than any of those listed above, one that scarcely has been considered in work on origins of language. “Functional flexibility” (15) is evidenced in infants when a single vocal category expresses positive, negative, and neutral emotional states on different occasions. This vocal capability can be implemented with any word or sentence in mature humans and is required for language because learning of culturally specific word meanings logically requires production of sounds that have no species-universal functions or meanings. We must be able to use any word or sentence with widely varying illocutionary forces (16), a fact that implies the ability to express a positive state as seen in joy or celebration, a negative state as in complaint, or an emotionally neutral state as can occur in factual description, all with the very same lexical or syntactic content. Thus, the sentence “the airplane is arriving” (or the single word “airplane”) can be used in these and many other ways: (*i*) in celebration of the landing of a flight, (*ii*) in complaint about the tardiness of a flight, or (*iii*) in mere description of a landing. The syntax and semantic content (or meaning) is the same in all three communicative acts, but the illocutionary content (or use) is vastly different. This sort of flexibility of use is not a dispensable side issue in language capability but rather is a fundamental requirement of all normal human communication.

Functional flexibility in vocalizations of infants as presented here implies that change in facial affect associated with any infant vocal category should correspond to reliably observable and predictable change in (*i*) the corresponding infant communicative act (e.g., the illocutionary force “complaint” should be associated with utterances showing negative facial affect, whereas “exultation” should be associated with utterances showing positive affect) and (*ii*) caregiver action in response to the infant social act (e.g., the perlocutionary effect of feeding an infant should occur with high likelihood when a negative vocalization is interpreted by a caregiver as indicating hunger,

Author contributions: D.K.O. and E.H.B. designed research; D.K.O., E.H.B., H.L.R., A.S.W., and L.C. performed research; D.K.O., E.H.B., L.C., and R.B. analyzed data; and D.K.O., E.H.B., H.L.R., A.S.W., and R.B. wrote the paper.

Conflict of interest statement: The authors declare no conflict of interest (such as defined by PNAS policy).

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: koller@memphis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1300337110/-DCSupplemental.

and continuation of protoconversation should occur when an infant vocalizes in a positive affiliative manner). The distinction between illocution and perlocution, drawn from Austin (16) and long used to advantage in descriptions of human infant communication (8, 17, 18), corresponds to a sender/receiver distinction in literature on animal communication, where it is emphasized that senders and receivers do not always have the same interests (19). For details on our use of the terms “illocution” and “perlocution” and our reasoning about their relation to the sender/receiver distinction, see *SI Appendix, Supporting Background: Affect and context in the judgment of function in vocalizations*.

It has been argued that, on balance, both senders and receivers must benefit from signal transmission for stable signals to evolve (20). In accord with this reasoning, illocutionary acts of senders on balance must predict perlocutionary effects in receivers, and the overall effect of signaling must be beneficial to both parties. This reasoning implies, as indicated above, that particular facial affect expressions through infant vocal categories should (if they have consistent pragmatic functions) correspond to particular classes of illocutionary forces and perlocutionary effects. An extensive literature in human caregiver–infant interaction (21–27), especially in the field of attachment (28–32), suggests that precisely this sort of correspondence between perceived affect and functions of infant communications occurs in real parent–infant interaction. Infants express affect, and parents interpret that affect functionally, responding with encouragement and/or continuation of pleasurable interaction when an infant interacts with positive affect but responding with physical attempts to change the situation by comforting or distracting and/or talking about a possible change in the situation when an infant expresses negative affect. The caregiver also responds to the infant with her own emotional expressions and affect, and thus such interaction has been thought to constitute a mutual emotional regulatory system (27, 33) with benefits to both parties in health and well-being.

The flexibility of affect expression and associated functions for linguistic units such as words stands in contrast with patterns reported in some of the literature on nonhuman primate vocalization where each vocal type or “call” has been portrayed as having a consistent function (*SI Appendix, Supporting Background*). This idea is rooted in the classical ethologists’ notion of “fixed signals” (34, 35), with each call seen as naturally selected to express a particular emotional or arousal state, implying transmission of a corresponding illocutionary force along with predictable perlocutionary reactions from listeners. Given this view, one would not expect individual nonhuman primate calls to change valence from positive to neutral to negative on different occasions of use.

In recent years, however, research has suggested that animal calls may be more functionally flexible than the classical ethologists imagined. For example, across development from infancy, chimpanzee grunts have been shown to occur in increasingly variable “contexts” that may suggest variable functions (36, 37). Additional sounds in chimpanzees and other primates later in life also have been shown to be used in a variety of contexts (38). However, the research has not yet shown that such variations extend to the sort of functional flexibility considered here, including variation in affect expression for a single call from positive to neutral to negative, nor has research illustrated that such variable affect in particular nonhuman primate calls corresponds to predictable variations in illocutionary force or perlocutionary outcome. By providing a framework to evaluate quantitatively the possibility that functional vocal flexibility occurs in infants, the present work may also provide a basis for future research in nonhuman primates to quantify important similarities or differences between humans and nonhumans on this critical feature of linguistic communication and its development.

Considerable research has suggested that human infant cry and laughter are closely related in terms of form, function, and brain control mechanisms to species-specific calls of nonhuman primates that may in fact be homologous with cry and laughter in humans (39–47). The evidence reported here, however, shows that in the “protophones,” the infant vocal types that seem to be

less related to our primate heritage and arguably more related to speech, emotional expression can change from utterance to utterance. The evidence also shows that these changes in affect correspond predictably to functional changes in classes of illocutionary force and of perlocutionary effect of infant utterances.

The protophones considered in the present research are squeals, vowel-like sounds (hereafter “vocants”), and growls (*SI Appendix, Supporting Background, Characteristics of the protophones*). These protophones require phonation (i.e., voicing) with or without supraglottal articulation; hence they may or may not contain distinguishable consonants and vowels. Protophones occurring before canonical babbling cannot be transcribed sensibly in the International Phonetic Alphabet (48), because they generally do not contain well-formed and distinguishable consonants and vowels. Precanonical protophone description instead requires ethological categorization (similar to that used in assessing nonhuman vocalizations) with intuitive listening, categorizations that can be supported by acoustic analysis (*SI Appendix, Supporting Fig. 1*), and statistical modeling (49). The protophone types appear to emerge during active infant vocal exploration (12), and this exploration may account for the fact that the categories are fuzzy, just as hand and arm movements constitute fuzzy categories with substantially variable trajectories during infant development of reaching and grasping (50–52). Despite this fuzziness, key protophones (including squeals, vocants, and growls) are reported spontaneously by parents and are consistently recognized by ethologically oriented researchers of infant vocalization (10, 53–55).

The existence of functional flexibility in human infant vocalization as well as its significance as a foundation for language has been largely ignored in the past because research on vocal development has focused heavily on discerning meaning in infant expression and thus on determining consistent rather than flexible functions (especially expressions of emotional state) for the seemingly disorganized vocalizations of infancy (56, 57). There previously has been no direct comparison illustrating the functional flexibility of protophones versus the functional fixedness of cry and laughter. This is a key gap because the demonstration of this difference could illustrate straightforwardly the very early emergence of a vocal capability present in and required for language and not yet reported in other primates.

Results

We analyzed longitudinal recordings of vocal interactions and play for nine infants at three ages in the first year. Acoustic analysts located 6,995 utterances, which were then coded for vocal type (cry, laugh, squeal, vocant, and growl) based on audio presentation only and for facial affect (positive, neutral, and negative) based on video only (*SI Appendix, Supporting Methods* and *SI Appendix, Supporting Table 1*). Fig. 1 illustrates that all the protophones showed predominant neutrality in facial affect (mean = 64%), suggesting that, in protophones, infants possess the requirement of language for flexible detachment of vocalization from particular emotions. Cries and laughs, in contrast, were rarely judged to be neutral in facial affect (mean = 4%). The figure also shows that cries were overwhelmingly deemed negative, whereas laughs were overwhelmingly deemed positive. In contrast, all the protophones showed both considerable numbers of utterances with positive affect and considerable numbers with negative affect. For examples of infant protophones illustrating functional flexibility, see *SI Appendix, Supporting Methods* and the audio/video examples (*Movies S1–S19*).

Fig 2 illustrates that the distribution of protophone affect differed starkly from cry and laugh affect in six ways: Protophones showed (i) more positivity than cry but (ii) less than laugh (Fig 2A), (iii) more neutrality than either cry or (iv) laugh (Fig 2B), and (v) less negativity than cry but (vi) more than laugh (Fig 2C). These six patterns were seen at all ages and in all infants (*SI Appendix, Supporting Figs. 6 and 7*) and were supported statistically by highly significant odds ratios (*SI Appendix, Supporting Tables 2 and 3*).

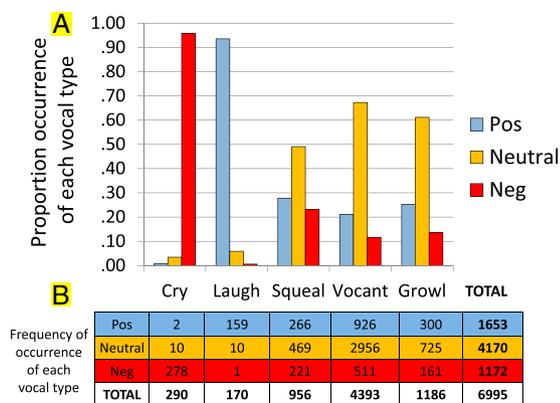


Fig. 1. Frequency and proportion of occurrence of each vocal type. Data on nine infants in the first year of life were collapsed across three observation periods. The emotional signals (cry and laughter) showed (A) proportions and (B) frequencies of occurrence with cries displaying overwhelmingly negative and laughs positive facial affect. In contrast, the protophones (squeals, vocants, and growls), presumed to be precursors to speech, were all used flexibly, showing primarily neutral facial affect and also presenting numerous cases of both positive and negative affect. For description of the vocal types, see *SI Appendix, Supporting Background*. Briefly, squeals are produced with very high pitch for the individual infant, vocants with mid-range pitch, and growls with harsh voice quality or low pitch. Positive, negative, and neutral facial affect correspond roughly to smiling/grinning, frowning/grimacing, and neither (see also *SI Appendix, Supporting Methods: Coding training and coding procedures for both vocal type and facial affect*).

As tabulated in Fig. 1, protophones occurred much more often than cry and laugh in our recordings, suggesting that vocalization even in the first months of human life is not characterized primarily by predetermined emotional expression but rather by exploratory vocal freedom and flexible expression, key foundations for language. Moreover, the most commonly occurring vocal type (vocant) was the one most commonly judged to be neutral in facial affect, suggesting again that these sounds reveal an emergent foundation for language—the possibility that vocalizations can be detached from any particular emotional state.

The data for the 10- to 12-mo-olds showed that infants often superimposed features of vocal quality corresponding to the three protophones upon the more mature vocal patterns, including canonical babbling; thus well-formed syllables were produced in the oldest infants with squeal, vocant, or growl-like

phonation, and these more complex vocal patterns continued to be expressed with a full range of facial affect.

Fig. 3 *A* and *B* demonstrate that facial affect expressed in protophones corresponded predictably to illocutionary force judgments based on simultaneous audio/video presentation. Protophones categorized as showing pursuit of protoconversation with the caregiver through comfortable vocal responsiveness, imitation, exultation, or initiation of interaction (the “Converse” group of illocutionary functions) were strongly associated with positive facial affect, whereas protophones interpreted as complaints or pleas for help (the “Complain” illocutionary functions) corresponded systematically to negative facial affect. Fig. 3 *C* and *D* similarly indicate that facial affect of protophones yielded predictable perlocutionary outcomes also judged based on simultaneous audio/video presentation. Thus, positive affect corresponded to caregiver encouragement of pleasurable vocal interaction through imitation, celebration, smiling, and so on (the “Encourage” perlocutionary group), whereas negative affect yielded explicit spoken assessment by the caregiver of a possible change in the physical situation or attempts to change the situation or the interaction (the “Change” perlocutionary group). All the relations between affect and both illocutionary and perlocutionary outcomes for Fig. 3 were supported by highly significant odds ratios ($P < 0.001$), and all the protophones showed similar patterns of relation between particular affect types and classes of both illocutionary forces and perlocutionary effects. Additional data on affect and its relation to function are provided in *SI Appendix, Results: The role of affect expression in the functional interpretation of infant protophones* and *SI Appendix, Supporting Figs. 9 and 10*. Possible coder bias in the assessment of the relation between facial affect and perlocutionary effect could be predicted if judgments of perlocution were driven by attention to both video and audio of both infant and parent. To determine if such bias could have been the source of the strong relation between facial affect and perlocutionary effects, we compared results from two coders, one of whom judged perlocutionary effects based on both audio and video of the caregiver and infant, and the other of whom made judgments based on caregiver utterances in audio alone, with no video and no child voice included. Results from both coders (*SI Appendix, Supporting Fig. 10*) showed, again with highly significant odds ratios, that perlocutionary effect judgments corresponded as predicted to facial affect judgments.

The consistency of the six patterns of difference in affect expression across the protophones should not be taken to mean that systematic patterns of protophone affect expression were entirely absent. Functional flexibility, we reason, does not correspond to random action but rather to the potential for systematic, adaptive

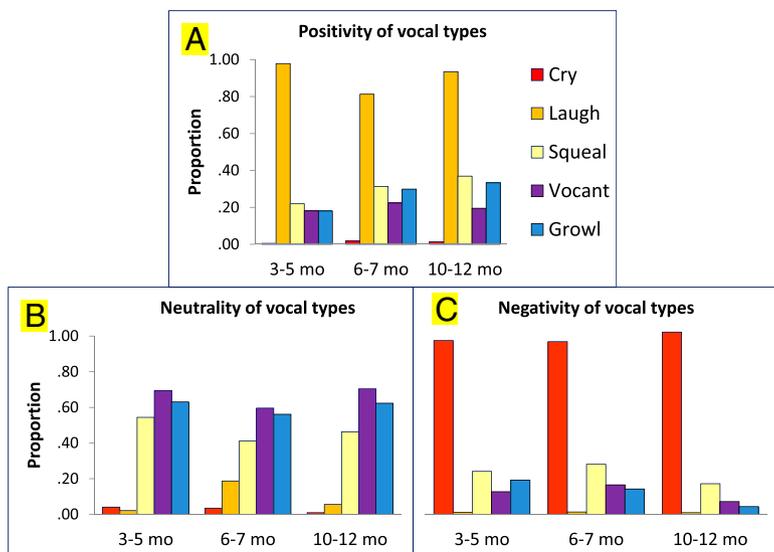
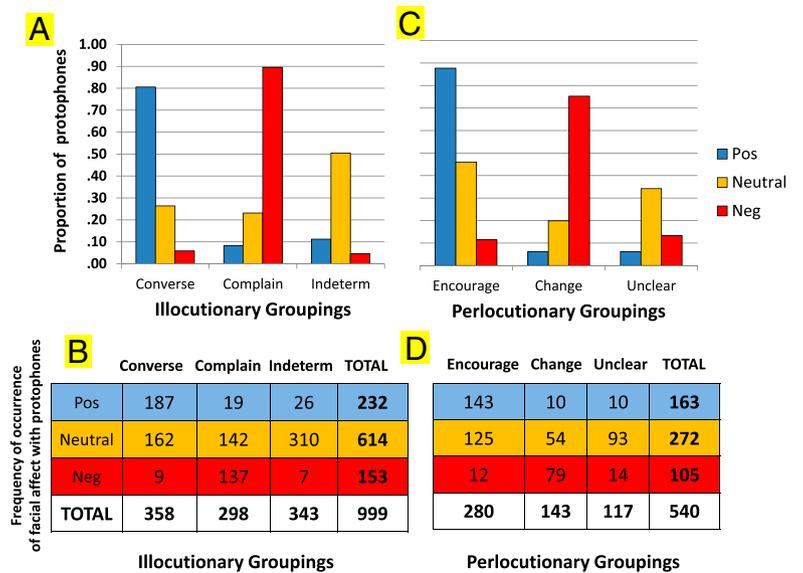


Fig. 2. Quantitative illustration of the distinction in functional flexibility between protophones and cry/laugh. Key patterns for protophones at all infant ages showed flexibility, as is required in speech, whereas cry and laugh showed consistent affect expression, as occurs in affectively charged vocalizations of other primates. (A) Positivity in facial affect expression: Protophones showed (i) far more positivity than cry but (ii) far less than laugh. (B) Neutrality in facial affect expression: Protophones were (iii) far more neutral than either cry or (iv) laugh. (C) Negativity in facial affect expression: Protophones were (v) far less negative than cry and were (vi) far more negative than laugh. Thus, the results (supported strongly by odds ratios; see *SI Appendix, Supporting Results: Odds ratio analyses* and *SI Appendix, Supporting Tables 2–3*) illustrate that in the first year, human infant protophones show the kind of flexibility in affect expression that is required for speech; such flexibility has not been reported as yet in nonhuman primate vocalizations at any age.

Fig. 3. Effects of facial affect on illocutionary force and perlocutionary effect of infant protophones. (A) Distribution of protophones with varying facial affect across three illocutionary groupings by proportion. (i) More than 80% of protophones with positive facial affect were coded as having illocutionary forces supporting protoconversation with the caregiver (e.g., initiating conversation, continuing conversation, expressing joy or exultation, imitating, and so on), whereas (ii) protophones with neutral affect most often were coded as indeterminate (e.g., they were not directed to the caregiver, often being interpreted as vocal play), and (iii) those with negative affect were coded 90% as complaints or pleas for help. (B) Distribution of protophones with varying facial affect across three illocutionary groupings by frequency of occurrence. (C) Distribution of protophones with varying facial affect across three perlocutionary groupings by proportion. (i) The vast majority (88%) of protophones with positive facial affect produced caregiver reactions encouraging protoconversation (e.g., calling to the infant, continuing conversation, praising, expressing joyful surprise, and so on), whereas (ii) those with neutral affect yielded more than twice as many encouragements as attempts to change the situation, along with many responses coded as ambiguous, and (iii) 75% of those with negative affect were responded to by explicit talk about what might be wrong with the baby or by attempts to change the situation or interaction (picking the baby up, attempting to distract him/her, soothing, and so on). (D) Distribution of protophones with varying facial affect across three perlocutionary groupings in frequency of occurrence. For details see *SI Appendix, Supporting Methods: Illocutionary force coding and perlocutionary force coding* and *SI Appendix, Supporting Results: The role of affect expression in the functional interpretation of infant protophones*, with *SI Appendix Supporting Figs. 9 and 10*.

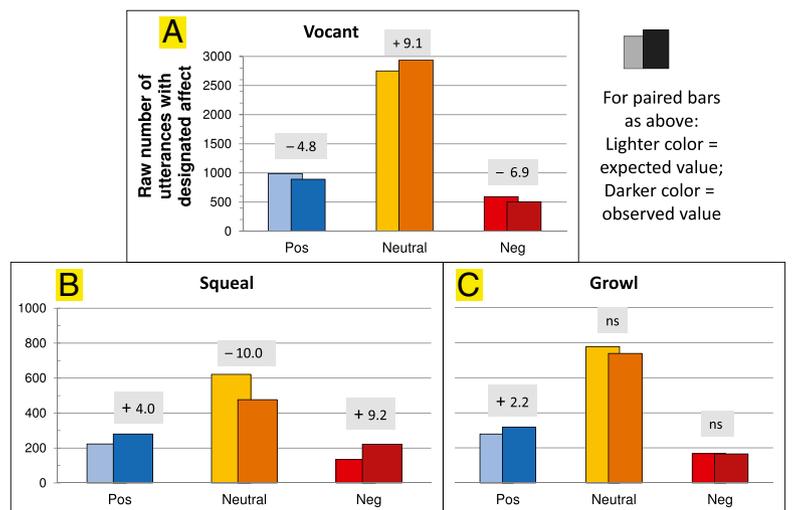


patterning, as is required in speech. The three protophones, although far more affectively flexible than cry and laugh, each displayed consistent patterns in affect expression. Fig. 4 compares the observed positive, neutral, and negative counts for each protophone with counts expected by an independence model assuming identical distribution of facial affect types across protophones. The independence model failed to fit; $\chi^2(4, n = 6,535) = 148, P < 0.001$. Adjusted residuals, approximating a unit normal distribution, revealed that the observed frequencies often differed by several SDs from the frequencies expected by the independence model, and the deviations differed notably across protophones. For example, vocants were 9.1 SD more likely than expected to be facially neutral, whereas squeals were 10.0 SD less likely than expected to be neutral. Both squeals and growls showed more than expected

positivity in facial affect, and squeals also showed more than expected negativity. Thus, squeals and growls were more likely than vocants to be used in expressing extremes of emotion, even though all three protophone types were used in substantial numbers to express both positivity and negativity.

Despite the systematicity of affect expression in protophones for data aggregated across infants (Fig. 4), contingency-table analysis revealed that the patterns varied by individual. Fig. 5 illustrates nine unique patterns, thus revealing an additional kind of functional flexibility—systematic individual variation in affect expression for protophones. Cry and laugh, on the other hand, showed no such individual variation: all infants displayed strong positivity in laugh and negativity in cry (*SI Appendix, Supporting Results: Additional contingency table analyses*,

Fig. 4. Observed and expected counts for each facial affect type, separately by protophone. Comparison of observed and expected counts, aggregated across observations of nine infants during the first year of life, illustrates that functional flexibility does not indicate lack of differentiation but instead indicates systematic flexibility in affect expression by infants across the three protophones, the presumed precursors to speech. Darker right-hand bars indicate observed counts of positive, neutral, and negative facial affect for each protophone. Lighter left-hand bars indicate counts expected if facial affect were independent of protophone type. Insets with a light gray background show adjusted residual values (some positive, some negative) corresponding to counts significantly greater than or less than expected ($P < 0.05$). (A) Compared with the counts generated by the independence model, vocants were significantly more likely to be neutral (9.1 SD more than expected) and were significantly less likely to be either positive or negative. (B) In contrast, squeals were more likely than expected to be both positive and negative and were less likely to be neutral, and (C) growls were more likely to be positive than expected by chance. These outcomes illustrate that although protophones manifest functional flexibility (all were predominantly neutral, and all also showed substantial numbers of cases of positive and negative expression), they were used affectively in systematically different ways with respect to each other. We reason that the ability to produce the same vocalization with differing affective character is a necessary foundation for language, but a random distribution of facial affect with respect to vocalizations would constitute a restriction on the very flexibility that is needed for adaptation of vocal expression to specific contexts and communicative needs. Thus, the systematic patterns in the data are consistent with the idea that infants control affective expression in protophones rather than producing facial affect in random association with the protophones.



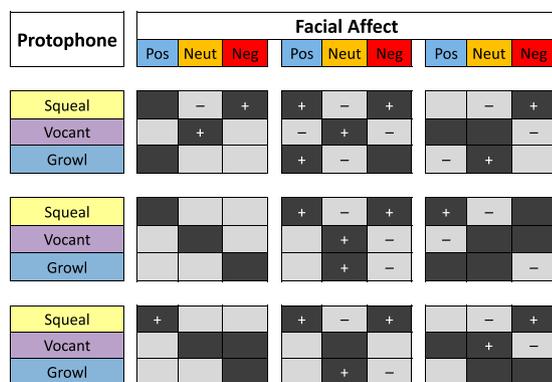


Fig. 5. Individual differences in facial affect expression through the protophones. Log-linear analysis of 3×3 tables for frequency of occurrence of protophones by facial affect (see text and *SI Appendix, Supporting Results: Log-linear analyses*) revealed significant differences among the nine infants. At the same time, log-linear analysis of 2×3 tables for cry and laugh revealed no such individual differences (because cries were negative and laughs positive for all infants who produced them; see *SI Appendix, Supporting Table 4*). This figure provides a quantitative illustration of individual differences in patterns of functional flexibility across infants for the protophones based on Contingency Table Analysis. Each square represents one of the nine infants. Cells within squares represent associations of protophones with facial affect types. Darker cells indicate positive adjusted residuals (observed counts greater than expected); a plus indicates an adjusted residual greater than +1.96. Lighter cells indicate negative adjusted residuals (observed counts less than expected); a minus indicates an adjusted residual less than -1.96. The figure indicates that each of the nine infants showed a unique pattern of adjusted residuals. These significant individual differences suggest that the functional flexibility of infants is not the result of an innate tendency specifying use of protophones in terms of affect (as appears to occur with cry and laugh) but rather that infants possess an inclination to explore vocalization in protophones and to be expressive with them, each infant thus developing a personalized path toward a capacity for speech (see also *SI Appendix, Supporting Results: Additional contingency table analyses*).

especially *SI Appendix, Supporting Table 4*). Log-linear analysis confirmed statistically significant individual differences for protophones but not for cry and laugh (*SI Appendix, Supporting Results: Log-linear analyses*, and *SI Appendix, Supporting Tables 5–7*), illustrating again the affect flexibility of protophones in sharp contrast with the affect rigidity of cry and laugh.

Perspective on protophones occurring in different contexts is supplied in *SI Appendix, Supporting Results: Robustness of functional flexibility of protophones across contexts*, where *SI Appendix, Supporting Fig. 11* illustrates that vocalizations occurring both during and not during gaze toward another person showed a strong pattern of variation in affect, similar to that reported in Fig. 1. Similarly *SI Appendix, Supporting Fig. 12* provides illustrations that infants used variable facial affect with all the protophones in each of five situational contexts, including one in which infants were not engaged in interaction but instead were vocalizing while playing alone in the same room with the parent. Thus, all the protophones occurred with variable facial affect in all the contexts, suggesting considerable robustness of functional flexibility across interactive contexts.

To assess observer agreement, two reliability coders judged 21% of the data for both facial affect and vocal type. They followed the same protocol as in the master coding that was used in our data analysis (*SI Appendix, Supporting Methods: Observer agreement levels for both vocal type and facial affect*). Agreement between the reliability coders and the master based on audio-only judgments showed $\kappa = 0.60$ for protophone category (squeal, vocant, or growl) and $\kappa = 0.91$ for cry and laugh. The much higher agreement for cry and laugh provides an additional indication of their relative innateness and immutability, whereas the lower agreement for protophones is consistent with their interpretation as emergent categories resulting from active exploration of vocalization

and perhaps of individual variation. Facial affect agreement for the reliability coders observing only video showed $\kappa = 0.73$ with respect to the master coding.

Observer agreement regarding the six patterns of functional flexibility differentiating protophones from cry and laugh was confirmed by an analysis of the same sort displayed in Fig. 2. All six patterns applied both to reliability coders and to the master coding, with protophones showing massively greater flexibility of facial affect than cry and laugh (*SI Appendix, Supporting Fig. 8*), patterns again strongly supported by statistically significant odds ratios.

Discussion

All the results converge on stark differentiation of the functional flexibility of protophones from the fixedness of cry and laughter by age 3–4 mo in human infants (*SI Appendix, Supporting Fig. 7*). The evidence suggests that the early protophones have a special role in language development and evolution because they are the first sounds to be free of specific fixed functions and thus reveal the opening of a door to the flexibility required for language.

Because vocal flexibility is a logical requirement for even the most rudimentary speech development, the evolution of language may have required the evolution of vocal flexibility such as seen in these human infants at a very early stage among hominins. An intriguing question is whether the evolution of such vocal flexibility was one of the first steps in communicative differentiation of the hominin line from that of other primates, especially our closest relatives, chimpanzees and bonobos. Research often has tended to emphasize limitations in flexibility of vocal communication in nonhumans and thus to suggest that the type of functional flexibility reported here for human infants may be absent or present only to a more limited extent in nonhumans. However, the precise studies that would need to be done for quantitative comparison of nonhuman primate vocal flexibility with the kind of flexibility indicated by the present results for human infants have not yet been conducted. Usually primatologists assess function by reasoning from contexts of use. Nonhuman primates often use a particular call in substantially varying contexts (37, 38), but such variation still may be consistent with “a generalized function that transcends the different contexts” (58, p. 185). A key goal is to determine the generalized function or functions that each call may serve, both in terms of illocution and perlocution. However, direct functional assessment will be difficult to apply consistently across species because of major differences in lifestyles. An approach focusing on affect expression may offer a useful first step toward optimal cross-species comparisons.

Thus, far, we do not know whether or to what extent nonhuman primate vocalizations may show reversals of valence in affect expression (and corresponding reversals of function) on different occasions, i.e., the sorts of reversals we have documented here for the human infant. A recent review emphasizes that little effort has been devoted to multimodal description of vocal communication in nonhuman primates (59), and consequently the emotional valence of vocalizations often has been assessed (to the extent that it has been assessed at all) in terms of external context rather than in terms of a combination of vocal patterns, facial affect, and other behavioral expressions of the sender. To conduct comparative studies patterned after the present one, the optimal approach would seem to require a scale of emotional valence applicable, *mutatis mutandis*, to nonhuman primates as well as human infants. Significant groundwork has been laid recently with the development of a facial affect coding system for chimpanzee (60, 61) modeled after the Ekman scheme for human facial expression (62). An optimal approach also may require consideration of additional expressive modalities, including for example gesture, posture, and piloerection to assign comparable labels of emotional valence across species.

If future research shows that nonhuman primates (especially apes) are incapable of vocal functional flexibility or (as seems also possible) are significantly less capable than human infants, we shall have isolated a key communicative differentiation that may have been one of the first on the route that led to language in hominins

after the split from the groups that would become chimpanzees and bonobos. If, on the other hand, the research determines that there is substantial functional flexibility in at least some of the calls of nonhuman primates, we shall have helped to cast further light on the evolution of language and its grounding in the primate lineage.

Methods

All procedures were approved by the University of Memphis Institutional Review Board. Details of methods are found in the *S1 Appendix*. Topics covered are (i) infants and recordings; (ii) selection of data for the present study; (iii) coding software; (iv) utterance location for coding; (v) coding training and coding procedures for both vocal type and facial affect; (vi)

Definitions of vocal types and facial affect types in the study; (vii) positive, neutral, and negative affect as a proxy for function; (viii) illocutionary force coding; (ix) perlocutionary force coding; and (x) observer agreement levels for both vocal type and facial affect.

ACKNOWLEDGMENTS. We thank Yuna Jhang and Beau Franklin for help with coding. The research for this paper was funded by Grants R01 DC006099 and DC011027 from the National Institute on Deafness and Other Communication Disorders and by the Plough Foundation, which supports D.K.O.'s Chair of Excellence at the University of Memphis. Theoretical work underlying this article was funded in part by the Konrad Lorenz Institute for Evolution and Cognition Research, where D.K.O. is an external faculty member. A.S.W. was supported by a U.S. Dept. of Energy Computational Science Graduate Fellowship, DE-FG02-97ER25308.

- Chomsky N (1993) *Language and Thought* (Moyer Bell, Wakefield, RI).
- Christiansen MH, Dale R (2004) The role of learning and development in language evolution: A connectionist perspective. *The Evolution of Communication Systems: A Comparative Approach*, eds Oller DK, Griebel U (MIT Press, Cambridge, MA).
- Pinker S (1994) *The Language Instinct* (Harper Perennial, New York), pp 494.
- Niyogi P (2006) *The Computational Nature of Language Learning and Evolution* (MIT Press, Cambridge, MA).
- Bickerton D (1990) *Language and Species* (Univ of Chicago Press, Chicago).
- West-Eberhard M-J (2003) *Developmental Plasticity and Evolution* (Oxford Univ Press, New York).
- Hall BK (1992) *Evolutionary Developmental Biology* (Chapman and Hall, London).
- Bates E, Benigni L, Bretherton I, Camaioni L, Volterra V (1979) *The Emergence of Symbols: Cognition and Communication in Infancy* (Academic, New York).
- Tomasello M (2008) *Origins of Human Communication* (MIT Press, Cambridge, MA).
- Oller DK (1980). The emergence of the sounds of speech in infancy. *Child Phonology*, Vol 1: Production, eds Yeni-Komshian G, Kavanagh J, Ferguson C (Academic, New York), pp 93–112.
- Stoel-Gammon C (2011) Relationships between lexical and phonological development in young children. *J Child Lang* 38(1):1–34.
- Oller DK (2000) *The Emergence of the Speech Capacity* (Lawrence Erlbaum Associates, Mahwah, NJ), pp 428.
- Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing intentions: The origins of cultural cognition. *Behav Brain Sci* 28(5):675–691, discussion 691–735.
- Locke JL, Bogin B (2006) Language and life history: A new perspective on the evolution and development of linguistic communication. *Behav Brain Sci* 29:259–325.
- Griebel U, Oller DK (2008) Evolutionary forces favoring contextual flexibility. *Evolution of Communicative Flexibility: Complexity, Creativity and Adaptability in Human and Animal Communication*, eds Oller DK, Griebel U (MIT Press, Cambridge, MA), pp 9–40.
- Austin JL (1962) *How to Do Things with Words* (Oxford Univ. Press, London).
- Dore J (1975) Holophrases, speech acts and language universals. *J Child Lang* 2(1): 21–40.
- Bates E (1976) *Language in Context* (Academic, New York).
- Owings DH, Morton ES (1998) *Animal Vocal Communication* (Cambridge Univ Press, Cambridge).
- Maynard Smith J, Harper D (2003) *Animal Signals* (Oxford Univ Press, Oxford).
- Bakeman R, Adamson LB (1984) Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Dev* 55(4):1278–1289.
- Beebe B, Alson D, Jaffe J, Feldstein S, Crown C (1988) Vocal congruence in mother-infant play. *J Psycholinguist Res* 17(3):245–259.
- Cohn JF, Tronick EZ (1987) Mother-infant face-to-face interaction: The sequence of dyadic states at 3, 6, and 9 months. *Dev Psychol* 23(1):68–77.
- Jaffe J, Beatrice B, Stanley F, Crown CL, Jasnow MD (2001) *Rhythms of Dialogue in Infancy: Coordinated Timing in Development* (Univ of Chicago Press, Chicago).
- Lester BM, Hoffman J, Brazelton TB (1985) The rhythmic structure of mother-infant interaction in term and preterm infants. *Child Dev* 56(1):15–27.
- Stern DN (1974) *mother and infant at play: The dyadic interaction involving facial, vocal, and gaze behaviors. The Effect of the Infant on Its Caregiver*, eds Lewis M, Rosenblum LA (Wiley, New York), pp 187–213.
- Tronick ED, Als H, Brazelton TB (1977) Mutuality in mother-infant interaction. *J Commun* 27(2):74–79.
- Bowlby J (1969) *Attachment and Loss* (Basic Books, New York).
- Ainsworth MDS (1969) Object relations, dependency, and attachment: A theoretical review of the infant-mother relationship. *Child Dev* 40(4):969–1025.
- Feldman R (2007) Parent-infant synchrony: Biological foundations and developmental outcomes. *Curr Dir Psychol Sci* 16(6):340–345.
- Forbes EE, Cohn JF, Allen NB, Lewinsohn PM (2004) Infant affect during parent-infant interaction at 3 and 6 months: Differences between mothers and fathers and influence of parent history of depression. *Infancy* 5(1):61–84.
- Schore AN (2001) Effects of a secure attachment relationship on right brain development, affect regulation, and infant mental health. *Infant Ment Health J* 22 (1–2):7–66.
- Feldman R, Greenbaum CW, Yirmiya N (1999) Mother-infant affect synchrony as an antecedent of the emergence of self-control. *Dev Psychol* 35(1):223–231.
- Lorenz K (1951) Ausdrucksbewegungen höherer Tiere. *Naturwissenschaften* 38(5): 113–116.
- Tinbergen N (1951) *The Study of Instinct* (Oxford Univ Press, Oxford).
- Laporte MNC, Zuberbühler K (2010) Vocal greeting behaviour in wild chimpanzee females. *Anim Behav* 80(3):467–473.
- Laporte MNC, Zuberbühler K (2011) The development of a greeting signal in wild chimpanzees. *Dev Sci* 14(5):1220–1234.
- Crockford C, Boesch C (2005) Call combinations in wild chimpanzees. *Behaviour* 142(4):397–421.
- Jürgens U (1995) Neuronal control of vocal production in non-human and human primates. *Current Topics in Primate Vocal Communication*, eds Zimmerman E, Newman JD, Jürgens U (Plenum, New York).
- Winter P, Handley P, Ploog DW, Schott D (1973) Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. *Behaviour* 47(3):230–239.
- Ploog DW (1992) *The Evolution of Vocal Communication. Nonverbal Vocal Communication*, eds Papoušek H, Jürgens U, Papoušek M (Cambridge Univ Press, New York), pp 6–30.
- Owren MJ, Amoss RT, Rendall D (2011) Two organizing principles of vocal production: Implications for nonhuman and human primates. *Am J Primatol* 73(6):530–544.
- Ross MD, Owren MJ, Zimmermann E (2010) The evolution of laughter in great apes and humans. *Commun Integr Biol* 3(2):191–194.
- Dunbar RIM (2004) Language, music and laughter in evolutionary perspective. *The Evolution of Communication Systems: A Comparative Approach*, eds Oller DK, Griebel U (MIT Press, Cambridge, MA), pp 257–274.
- Kipper S, Todt D (2003) Dynamic-acoustic variation causes differences in evaluations of laughter. *Percept Mot Skills* 96(3 Pt 1):799–809.
- Newman JD (1985) The infant cry of primates: An evolutionary perspective. *Infant Crying: Theoretical and Research Perspectives*, eds Lester BM, Boukydis CFZ (Plenum, New York), pp 307–323.
- Vettin J, Todt D (2005) Human laughter, social play, and play vocalizations of non-human primates: An evolutionary approach. *Behaviour* 142(2):217–240.
- Lynip AW (1951) The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant. *Genet Psychol Monogr* 44(2):221–262.
- Warlaumont AS, Oller DK, Buder EH, Dale R, Kozma R (2010) Data-driven automated acoustic analysis of human infant vocalizations using neural network tools. *J Acoust Soc Am* 127(4):2563–2577.
- Smith LB, Gasser M (2005) The development of embodied cognition: Six lessons from babies. *Artif Life* 11(1–2):13–29.
- Thelen E (1981) Rhythmic behavior in infancy: An ethological perspective. *Dev Psychol* 17(3):237–257.
- Thelen E (1995) Motor development. A new synthesis. *Am Psychol* 50(2):79–95.
- Koopmans-van Beinum FJ, van der Stelt JM (1986) Early stages in the development of speech movements. *Precursors of Early Speech*, eds Lindblom B, Zetterstrom R (Stockton, New York), pp 37–50.
- Roug L, Landberg I, Lundberg LJ (1989) Phonetic development in early infancy: A study of four Swedish children during the first eighteen months of life. *J Child Lang* 16(1):19–40.
- Stark RE, Bernstein LE, Demorest ME (1993) Vocal communication in the first 18 months of life. *J Speech Hear Res* 36(3):548–558.
- Papaeliou C, Minadakis G, Cavouras D (2002) Acoustic patterns of infant vocalizations expressing emotions and communicative functions. *J Speech Lang Hear Res* 45(2):311–317.
- Scheiner E, Hammerschmidt K, Jürgens U, Zwirner P (2002) Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *J Voice* 16(4):509–529.
- Notman H, Rendall D (2005) Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Anim Behav* 70(1):177–190.
- Slocombe KE, Waller B, Liebal K (2011) The language void: The need for multimodality in primate communication research. *Anim Behav* 81(5):919–924.
- Parr LA, Waller BM, Heintz M (2008) Facial expression categorization by chimpanzees using standardized stimuli. *Emotion* 8(2):216–231.
- Parr LA, Waller BM, Vick SJ (2007) New developments in understanding emotional facial signals in chimpanzees. *Curr Dir Psychol Sci* 16(3):117–122.
- Ekman P, Friesen W (1978) *The Facial Action Coding System* (Consulting Psychologists, Palo Alto, CA).