The function and evolution of child-directed communication

Johanna Schick[1,5¶], Caroline Fryns[2,5¶], Franziska Wegdell[1,5¶], Marion Laporte[7,8], Klaus Zuberbühler[2,4,5], Carel P. van Schaik[3,5,9&*], Simon W. Townsend[1,5,6&], Sabine Stoll[1,5&*]

1 Department of Comparative Language Science, University of Zurich, Zurich, Switzerland

2 Department of Comparative Cognition, University of Neuchatel, Neuchatel, Switzerland

3 Department of Anthropology, University of Zurich, Zurich, Switzerland

4 School of Psychology and Neuroscience, University of St. Andrews, St. Andrews, United Kingdom

5 Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, Zurich, Switzerland

6 Department of Psychology, University of Warwick, Warwick, United Kingdom

7 Histoire naturelle de l'Homme préhistorique, UMR 7194, PaleoFED, Muséum National d'Histoire Naturelle, Paris, France

8 Institut des Sciences du Calcul et des Données, Sorbonne Université, Paris, France

9 Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

* sabine.stoll@uzh.ch (SS); *vschaik@aim.uzh.ch (CPvS)

[¶]These authors contributed equally to this work.

[&]These authors also contributed equally to this work.

Abstract

Humans communicate with small children in unusual and highly conspicuous ways (child-directed communication (CDC)), which enhance social bonding and facilitate language acquisition. CDC-like inputs are also reported for some vocally learning animals, suggesting similar functions in facilitating communicative competence. However, adult great apes, our closest-living relatives, rarely signal to their infants, implicating communication surrounding the infant as the main input for infant great apes and early humans. Given cross-cultural variation in the amount and structure of CDC, we suggest that child-surrounding communication provides essential compensatory input when CDC is less prevalent— a paramount topic for future studies.

Introduction

Human languages exhibit enormous variation at all linguistic levels, ranging from phonemes, the smallest meaning-distinguishing units, to morphemes, the smallest meaning-bearing units, to words, higher-level constructions, and rules of combination. Few, if any, of these features are under strong genetic control. As a consequence, linguistic units must be learned from scratch by every maturing individual: a process that, while often described as “effortless” [1], in fact takes many thousands of hours of exposure over multiple years. Inevitably, the communicative environment must provide the input required for learning a native language.

One prominent source of this input is a special speech register used by caregivers to address infants and young children, frequently referred to as babytalk, motherese, parentese and, more recently, infant-directed or child-directed speech [2]. In this Essay, we use a more neutral term child-directed communication (CDC, see Box 1) since there is lack of agreement of what constitutes infancy in humans, and moreover, the input is modality independent (i.e., it is also encountered in sign languages [3,4]). Such cross-modal prevalence has even been argued to support the notion that CDC is an automatic and potentially species-wide trait [5]. Both in signed and spoken languages, CDC includes other multi-modal features such as more exaggerated facial expressions [6], modified gestures [7] and motions in general, with the latter known as motionese [8].

Box 1. Definitions of key terms

Child-directed communication (CDC): All communication specifically directed at children, in which the properties and structure of the signal often change in predictable ways, e.g. higher pitch, more exaggerated gestures and more repetition. CDC supports language learning in children [2,9].

Child-surrounding communication (CSC): All communication that is perceptible to the child but not directed at them.

Immature-directed communication: All communication specifically directed at the immature animal, as indicated by the vocalizations or gestures being accompanied by body or head orientation towards the immature animal, as well as a change in structural or acoustic features, for instance more repetition.

Natural pedagogy: The specific aspects of human communication that allow and facilitate the transfer of generic knowledge to novices [10].

9-month revolution: A large set of cognitive and socio-cognitive skills that human infants typically develop at around 9-12 months of age. Within this skill set they develop the ability to use gaze-following, social referencing, pointing, joint attention and imitation to join the adult’s attentional focus [11]. They also become able to interpret adults’ gestures as intentional acts [12].

Vocal learning: describes vocal production learning, which is traditionally defined as the production of novel vocalizations as a result of learning from an acoustic signal [13]. Today, many dimensions and degrees of vocal production learning are acknowledged [14]. Only few animal species are known to be capable of vocal production learning (e.g. songbirds, hummingbirds, cetaceans and pinnipeds). In contrast to vocal production learning stand usage and comprehension learning, which are more common in a wide variety of species [15]. Usage learning is defined as learning to produce a signal in a new context as a result of acoustic experience. Comprehension learning is defined as learning a new meaning of a signal as a result of experience [13].

A second and much less researched source of input is child-surrounding communication (CSC, Box 1), which includes all communication that is in perceptible proximity to, but not specifically directed towards the child. Typically, this involves two or more individuals engaged in some type of social interaction accompanied by a linguistic exchange. It may also include linguistic input from media sources (e.g. TV, radio), but it remains unclear which impact this type of input might have on the child's language development. CSC input is ubiquitous, and at least as omnipresent as CDC, yet we know much less about its functional role in language acquisition. The few available studies on CSC suggest that it has less impact than CDC on linguistic development in early ontogeny [16,17] .

The reliance on child-directed communication for the acquisition of communicative competence may be explained by three distinct evolutionary pathways (Fig 1). First, it might be shared with our closest living relatives, the great apes. If this is the case we can assume that it is a feature that was also present in early hominins (i.e. the "African Apes"; extant and extinct *Homo*, *Pan* and *Gorillini* genera). Second, it may be derived in humans, and perhaps be one of the drivers of the evolution of language, potentially as part of a wider change in cognitive architecture of early humans. This derived state can have arisen uniquely in our ancestors or, third, it can be fully or partially shared with other, distantly related taxa, in which case it arose via convergent evolution.

Fig 1. Evolutionary pathways of child-directed communication. A feature such as child-directed communication with the function of aiding the acquisition of communicative competence can be (1) ancestral: homologously derived among African great apes and thus also found in humans, (2) unique among the great apes but convergently shared analogously with other, more distantly related species, or (3) newly evolved within our own species. Red represents the presence of immature-directed communication features. Outline credits: Human - T. Michael Keesey; Chimpanzee: Jonathan Lawley; Bonobo: T. Michael Keesey; Gorilla: T. Michael Keesey (after Colin M.L. Burnett); Orangutan: Gareth Monger; Gibbon: Kai R. Caspar; Tamarin: Yan Wong and T.F. Zimmerman; Zebra Finch: Jim Bendon (photography) and T. Michael Keesey (vectorization); Bat: Yan Wong; Squamate: Ghedo and T. Michael Keesey; Feline: Margot Michaud; Equine: T. Michael Keesey; Cetacean: Scott Hartman; Falcon: Liftarn; Fish, macaque and baboon are uncredited. Link to creative commons license: https://creativecommons.org/licenses/by-sa/3.0/. Link to public domain license: https://creativecommons.org/publicdomain/zero/1.0/. Outlines were downloaded from http://www.http://phylopic.org/. The layout of the figure was achieved in R (version 4.1.2, R Development Core Team, 2012).

Current evidence suggests that in non-human primates in general (hereafter primates) the ability to produce species-specific vocalizations develops with relatively little environmental contribution, i.e., irrespective of auditory input [18-21]. Instead, input seems to have more of a role in guiding vocal usage and comprehension [22-25]. Nonetheless, at least some vocal production, flexibility does exist in primates, although mainly in terms of socially driven vocal accommodation [22,26-32]. Although this suggests a role for social input, how much of this is immature-directed communication (IDC) versus immature-surrounding communication remains unclear [33]. So far, the few studies that have assessed immature-directed vocalizations in great apes have yielded low rates (chimpanzees, *Pan troglodytes* [33]; bonobos, *Pan paniscus* [34]). A few studies have described vocalizations used by mothers in chimpanzees [35] and orangutans [36]. However, this directed communication does not display any of the features or functions of natural pedagogy. Overall, the current state-of-the-art suggests that immature-directed input has only a small impact on great ape vocal ontogeny, if any. The preliminary conclusion thus appears to be that most acoustic features of CDC are derived in humans. However, in the structural domain some precursors of CDC might exist in apes.

However, a striking exception is found in the gestural domain. Orangutans [37], chimpanzees [38] and bonobos [39] all use immature-directed gestures. Furthermore, one CDC-like feature, repetition, is found in gorilla [40] and chimpanzee gestures [41]. The use of specific gestures and their repetition rates by adult great apes towards immature individuals varies depending on the age and experience of the immature animal, as in humans, suggesting functional significance in the acquisition of communicative competence [40,41]. However, repetitions of gestures following lack of comprehension have also been described in adult orangutans [42]. In addition, bonobos modify communication signals according to recipient familiarity [43]. All of this suggests at least some shared cognitive features with humans. Evidently, more research is needed to assess whether immature-directed gestures can be considered the functional equivalent of CDC, especially in light of suggestions that at least part of the gestural repertoire are the result of innovations and therefore have to be learned [44].

If CDC is fully or at least partially derived in humans, this raises two important questions. First, which elements of the broad bundle of features that make up human CDC were already present in the last common ancestor? Identifying which elements were pre-existing (homologies: present in great apes), which are found in other animals (analogies: convergently evolved), and which are new and uniquely derived in our lineage would improve our understanding of how language acquisition evolved (Fig 1). Second, as IDC in primates in general appears to be rare, primates must acquire the learnt part of their communication from the communication that surrounds them, but is largely not directed at them. Has this originally predominant source of input remained significant in humans, or has CDC replaced it (Fig 2)?

Fig 2. Transition of child-surrounding to child-directed communication. The transition of the importance of use of child-surrounding communication (CSC) to child-directed communication (CDC). Darker color shows importance/presence and brighter color possible insignificance of CSC and CDC from early hominins to extant humans.

In this Essay, we aim to address these two questions. In the first section, we deconstruct CDC into its component parts and assess their proposed functions; we then ask for each of them whether comparable phenomena exist in non-human animals (hereafter animals). In the second section, we contrast CDC in humans with the lesser-studied CSC to shed light on the interplay between these two forms of input and their respective roles in language acquisition. Answers to these questions should not only improve our understanding of the development and acquisition of language but also its evolutionary progression.

The features and functions of CDC

CDC differs from adult-directed communication in a wide range of acoustic and structural features. This has been observed in numerous cultures and is widely considered a universal of human language [9, 45, 46]. Over the past few decades, a plethora of studies have shown that features of CDC (Table 1) support language acquisition by infants both in comprehension [47, 48] and production [49-51]. CDC is part of a more general package of child-directed behaviors that serve to pass on cultural knowledge and skills to the next generation, known as natural pedagogy [10] (Box 1). This active transmission process rests on an (arguably) uniquely human capacity, ostension, which underlies pointing and results in gaze following (often followed by joint attention on objects between caretaker and child [52] or a state of shared intentionality more broadly [11]), as well as child-directed speech [10]. In this Essay, we argue that CDC is a crucial part of this universal form of teaching. Such natural pedagogy is almost certainly derived relative to the non-human great apes (hereafter great apes) and potentially evolved in relation to the frequently highlighted shift in the breeding system from independent to more cooperative [53]. Although the child-development literature may seem to suggest that natural pedagogy is primarily aimed at preverbal infants and mainly geared toward teaching cultural knowledge, CDC is an obvious and essential part of natural pedagogy extending well beyond early infancy. In fact, one might hypothesize that CDC is a core feature enabling the transmission of language and as a consequence the evolution of such a complex communication system.

Table 1. Known features of child-directed communication (CDC).

Type of feature	Known feature of CDC	Proposed function	Reference
Acoustic	Pitch variability	Attention grabbing	[54]
Acoustic	Lengthening of vowels and pauses	Segmentation and discrimination of sounds	[55,56]
Acoustic	Extended vowel triangle	Sound discrimination	[57]
Acoustic	Clear articulation	Facilitate comprehension	[46,58]
Acoustic	Increased voice-onset time	Sound discrimination	[59]

Acoustic	Slower speaking rate	Facilitate comprehension, discrimination and segmentation	[54,60]
Structural	Frequent repetitions	Structural generalization of word/unit classes	[61,62]
Structural	Short utterances	Facilitate comprehension	[63,64]
Structural	Low type/token ratio	Facilitate comprehension	[65,66]
Structural	Simplified syntax and semantics	Facilitate comprehension	[63,65]
Structural	Frequent use of diminutives	Simplification of certain morphological aspects (language specific)	[67,68]
Structural	Frequent questions	Invite response, repetition, attention grabbing	[69,70]
Structural	Variation sets	Structural generalization of word classes	[71,72]
Structural	Scaffolding	Learning of word constructions	[73]

The first eight entries above the bold dividing line represent elements where a corresponding form could possibly be present in animal vocal communication.

Adults and older children use the bundle of acoustic and structural features of CDC in varying combinations when talking to infants and younger children (Table 1). For many of these features, there is evidence that they facilitate the child's language learning.

Regarding the prosodic and acoustic features of the speech, CDC involves the production of higher and more variable pitch [54], systematic lengthening of vowels and pauses [55,56,74] and an extended "vowel triangle" or vowel hyperarticulation [57,75]. Studies have shown that these prosodic modifications attract the child's attention [76] from an early age and that CDC is more salient to children than adult-directed communication and is actually preferred by them [60,77-79]. Indeed, neurobiological research has revealed that an infant's exposure to CDC in their first year of life results in a higher brain activation in their left and right temporal areas compared with adult-directed speech [80]. These prosodic modifications also elicit increased infant vocal responses during their prelinguistic phase [81], a form of active participation crucial to language acquisition [2]. Infants listening to CDC rather than adult-directed speech also show greater sensitivity to syllable and vowel discrimination [75,82]. Lastly, caregivers tend to use exaggerated prosody to mark new or relevant vocabulary [74,83,84]. These prosodic characteristics of CDC not only support the detection of word boundaries [85], but also word comprehension [48,86] and production [49]. In sum, acoustic alternations of the speech signal appear to accelerate various aspects of language acquisition (see [87] for a review), suggesting that CDC serves as an evolved teaching tool.

Regarding the structural features, CDC is characterized by short utterances [63,64], a low type/token ratio [65,66], which indicates that caregivers use a simplified vocabulary, and the use of many questions [69,70], diminutives [67,68] and repetitions [61,62]. One structural feature in particular is known to have a significant role in the acquisition of language: frequency effects. The more frequently an element occurs in the child's input, the faster it is expected to be learned [88,89]. Recent research has also shown that frequent repetitions are structured in CDC. Repetitions of constructions at the beginning of utterances (e.g., this is an X [62,90]) and discontinuous repetitions (e.g., I X you [91,92]) are ubiquitous and support the generalization of word classes, such as nouns and verbs [93]. In addition, repetitive structures or distribution of words surrounding specific verbs support the generalization of meaning [94], and the high number of repetitions found in CDC are positively correlated with word comprehension [95,96]. A specific form of repetitions frequently used in CDC is variation sets, successive utterances with partial self-repetitions produced by caregivers [71,72], which themselves are positively related to better linguistic outcomes in naturalistic longitudinal [97] and experimental settings [98]. These findings again support the hypothesis that CDC functions to accelerate language acquisition.

In addition to the prosodic and structural features of CDC, another important factor is the absolute amount of linguistic input children receive. A number of studies have indicated that the amount of CDC children experience is correlated with their later vocabulary development [16,99-102] and their word processing skills [101]. The quality (variety of words and syntactic structures) of CDC also impacts language development. Longitudinal studies have shown how input quality at an earlier stage of development predicts subsequent diversity and variance in language outcome at a later stage of development [103,104]. Quality and quantity may even have different roles during the child's language development. For example, a longitudinal study of vocabulary acquisition revealed that input quantity mattered most during the second year of development, whereas input quality was more important during the third year [50]. The child's ability to profit from different properties of CDC might therefore vary across development.

Most of the previously reviewed evidence is from children growing up in modern Western societies, characterized by child-rearing practices that are very different from what is typically seen in hunter-gatherer groups, our evolved and species-typical way of life [17]. In addition, there is substantial variation both within and across cultures in the amount of CDC that occurs and its features. Also important is that, in terms of sheer amount, there are linguistic communities in which children are only rarely directly addressed by their caregivers [105,106], suggesting that CDC is not essential for language acquisition, at least not as the main source of linguistic experience. A comparative study by Shneidman et al. [16] demonstrated that for 1-year-old children growing up in a Yucatec Mayan community the mean number of utterances a child encountered per hour amounted to approximately 400 utterances, with only 20% of it being directed to the child. The US group of 1-year olds that served as a comparison were exposed to approximately 900 utterances per hour, with more than 70% of these utterances being directed. More recent studies from non-WEIRD (Western, Educated, Industrialized, Rich and Democratic [107]) cultures confirmed that the amount of directed communication children are exposed to can vary strongly (e.g., Netherlands: 303 vs. Mozambique: 58 utterances of CDC/30 min [108]; Tseltal: 3.63 min of CDC/hour [109]; Tsimane: $>$1 min/daylight hour [17]; North American: 11.36 min of CDC/hour [110]), raising questions about the relevance of CDC as the critical source of language acquisition. So far the factors determining the amount of CDC are unclear. In particular the role of the child in the society might be crucial, i.e., whether a society adapts situations to the child or expects to the child to adapt to the situation [106,111].

Nonetheless, various studies revealed the presence of CDC features in non-WEIRD cultures (e.g. higher pitch [112]; slower speaking rate [113]; repetitions, diminutives and simpler syntax [114]). Overall the results suggest that both similarities (e.g. in pitch [113]) and differences [115] between WEIRD and non-WEIRD cultures do exist. However, not all CDC features can be found in every culture. In Quiché Mayan, for example, mothers do not seem to produce higher pitch when talking to their children, potentially because they must use this register when speaking to a person of higher status [116].

At this stage, it seems that the only universal characteristic of CDC is the presence of repetitive structural patterns in the input. Clearly, generalizations would be premature until more research reveals patterns linked to the social organization of a linguistic community. However, if one considers CDC as a toolkit, the main features of CDC (Table 1) presumably change gradually as the infant progresses to being a toddler and preschooler [117-120]. During the earliest stage before the 9-month revolution [12] (see Box 1), acoustic and structural features appear to be very prominent, whereas structural features seem to gain greater prominence at later stages (Table 1). Thus, initially the function of CDC may be to establish and strengthen the social bond with infants, direct attention [121], introduce turn-taking via proto-conversations [122], and scaffold the learning of the prosody, phonemes, morphemes and first words of the local language. After the 9-month revolution, once joint attention, intention reading, symbol recognition and rational imitation [11] have emerged, CDC may instead be geared more toward the learning of vocabulary and grammar.

A key next step in research would be to determine, for each culture, which features occur at what stage in development and in which combination, and how these tools interact. CDC might turn out to be heterogeneous across cultures. This variation might then be linked to the age at which children achieve adult-level competence in the various components of language.

The features and functions of immature-directed vocalizations in animals

To identify both the evolutionary roots and adaptive functions of CDC in humans we must examine similar phenomena in animals. We already noted that preliminary work on great apes suggests our common ancestor featured few, if any, of the elements of CDC as listed in Table 1, at least in the vocal domain. However, it must be stressed that this absence may simply reflect a lack of focused research effort rather than actual absence. But if it is confirmed, this would suggest that surrounding vocalizations provide the primary input for the learned part of the vocal development in great apes and that CDC originated *de novo* in the human lineage (Fig 1), presumably linked to the emergence of natural pedagogy, which may have preceded, and in fact facilitated, language evolution [53].

We now turn to possible convergent cases. First, we already discussed calls by great ape mothers, but they also occur in other primates [123,124], as well as in many non-primate species, where mothers call to their infants to retrieve them. Examples include domestic cats (*Felis silvestris catus* [125]), and ungulates such as domestic sheep (*Ovis aries* [126]), cattle (*Bos taurus* [127]), goitred gazelles (*Gazella subgutturosa* [128]) or saiga antelopes (*Saiga tatarica tatarica* [129]). Second, immature-directed calls may serve to aid recognition of the mother's voice, as in domestic cats [125], Mexican free-tailed bats (*Tadarida brasiliensis mexicana* [130]), fur seals (*Arctocephalus tropicalis* [131]), or domestic sheep [126]. These examples show that even if IDC exists in an animal species, it is unlikely that these cases are functionally equivalent to human CDC.

However, in a third category of species, we find immature-directed calls related to their capacity for vocal accommodation (small alterations of vocalizations as a result of experience [132]) and vocal learning (Box 1). Orcas (*Orcinus orca*) produce family-typical calls at higher rates after the birth of a calf [133]. Likewise, common marmosets (*Callithrix jacchus*), which show evidence of accommodation learning, and thus some level of vocal plasticity [134], modify call rates and repeat various different call types before and after birth of infants [135]. In agile gibbons (*Hylobates agilis*), duetting by mothers with inexperienced young has also been argued to represent IDC, serving to aid the acquisition of the species-specific song [136]. In these cases, the calls may serve to acquire the group's vocal signature.

Finally, some cases show suggestive parallels to human CDC. In cooperatively breeding marmosets, adults give contingent vocal feedback specifically to infants, which is suggested to impact vocal ontogeny since infants exposed to more of such calls by adults produce and properly use adult-like calls earlier [28,137], possibly owing to increased practice or because vocal feedback reduces stress [13]. This contingent vocal feedback may help infants acquire the underlying rules of dyadic vocal communication (i.e., turn-taking [138], but see [139]). Outside primates, in zebra finches, male tutors use a more stereotypic song when they are near immature birds [140]. In greater sac-winged bats (*Saccopteryx bilineata*), mothers adjust the pitch and timbre when they use immature-directed vocalizations [141].

Despite these parallels, no study has asked exactly which features of the vocalizations (Table 1) are essential and which functions they serve. It is therefore too early to conclude the common incidence of CDC-like functions of immature-directed vocalizations in either primate or non-primate species [28,40,140-142]. Systematic comparisons are needed to assess the extent of convergence and the determinants, but it remains plausible that IDC serves to facilitate the learning of vocal signatures (in accommodators) or call repertoires (in vocal learners sensu stricto), similar to the language-acquisition function of human CDC.

The function of CDC relative to CSC in humans

Although considerable attention has been paid to CDC and its structuring and function, comparatively less is known about the relative role of surrounding communication that children are exposed to (CSC). Indeed, in some linguistic communities surrounding communication is the primary source of input since adults rarely directly address infants (e.g. Kaluli and Samoan [106]; Yucatec Mayan [16], Tsimane [17]), at least in their first year of life. Despite these differences in input type, children still become competent native speakers [106,109,143,144]. This inevitably begs the question how important CDC actually is for speech development and suggests that CSC, though currently still under-researched, may have an equally important, perhaps compensatory role in facilitating language acquisition. In small-scale societies, which arguably represent the more typical human condition, children are continuously surrounded by individuals of all ages [145], suggesting that the amount and variation of CSC will be higher than in WEIRD societies. To date, the few studies that to our knowledge have quantitatively assessed this [17,109,146] have not revealed an effect of CSC on vocabulary development [16,101]. However, more work is needed to understand whether CSC supports the learning of other properties of language such as grammatical features.

To obtain a full understanding of how communicative competence develops in both humans and animals, it is critical to account for both sources of input — CDC and CSC — and the interplay between them. Are both CSC and CDC essential for proper language learning, or are they to some extent compensatory? If so, do the large amounts of CDC in WEIRD societies serve to compensate for the much lower quantity of CSC? In animals, immature-surrounding vocalizations might well be the predominant form of input, yet very little research has attempted to quantify their occurrence and assess their influence on the development of communicative competence. Filling this gap should be a high priority for research.

The question arises whether the relative amounts of CDC and CSC seen in humans are comparable to those found in great apes. The one study on chimpanzee infants suggests that immature-surrounding communicative events total approximately 15 gestures, 50 vocalizations and 3 gesture-call combinations per hour [147]. This is considerably more than what is known so far about the above mentioned low rate of immature-directed vocalizations. In all likelihood, therefore, immature-surrounding vocalizations were the most important source for the learnt part of the vocal system (usage and comprehension learning) in early hominins.

Conclusion and future directions

In human language learning, the amount and quality of CDC is one of the key facilitators of learning. But how the various features that make up CDC change with age, especially relative to the 9-month revolution, is not clear, and should be the target of future studies because they may vary in function from creating attachment, to establishing joint attention, to supporting specific details of language acquisition.

Despite its universality, research across and within cultures has shown enormous variation in a child's exposure to directed communication. Studies of a few non-WEIRD societies show much lower rates of CDC than found in the typical studies of WEIRD societies. This suggests that the amount of CDC children are exposed to in WEIRD societies might be atypical for the rest of the world and most of human history. Given the fact that all children learn the language of their culture, independent of culture-specific variation in input, the role of CSC for language learning might have been underestimated. The increased amount of CDC in WEIRD societies seems to result mainly in a refinement of skills, involving the size of the vocabulary and the construction inventory involved. This raises the question how CDC produces this refinement. Its impact may relate to the interactional situations in which it occurs. In these contexts joint attention is the key component that actually facilitates learning [52,148,149]. Such joint-attentional frames allow the reduction of interpretation space of form-meaning associations. Given the extreme cross-linguistic variability of CDC we must ask the questions of whether and how much CDC is really essential to language learning, whether CSC would do an equivalent job but just more slowly, or whether CDC is essential at particular stages only. Daylong recordings in naturalistic conditions are likely to provide answers to these questions.

To shed light on how CDC evolved, we examined research on our closest relatives, the great apes. So far very little directed input to infants has been documented. Concerning the features of human CDC (Table 1), few have been found in ape communication, except for repetition of gestures. Repetition is arguably the best predictor of language acquisition in human infants and children [88,89,150]. These findings suggest that short-term repetitive use of communicative acts is potentially an ancestral feature of CDC. We therefore propose more research is needed on structural repetition to complement the usual emphasis on acoustic features of CDC.

With regard to other animal species there is more evidence for immature-directed vocalizations in species that engage in vocal learning. This supports the idea that CDC in hominins arose to support the acquisition of highly culturally variable acoustic and structural features of language. However, much more systematic comparisons are needed, which should indicate which of the features characterizing human CDC are also found in these convergent cases. Obviously, more targeted work on great apes is a high priority, if only to see whether repetition is the only CDC-like feature present and why gestures appear to be the exception.

In sum, the current state of research suggests that most features of human CDC have evolved anew in our hominin ancestors. It serves to engage children in social interaction with caretakers and thus to facilitate language acquisition, and in later phases more explicitly in the acquisition of semantics and grammar. In other words, there is no doubt that CDC is an implicit teaching device. Doubt remains, however, whether it is the only facilitator.

References

1. Kuhl PK. Early language acquisition: cracking the speech code. Nat Rev Neurosci. 2004;5(11):831-43. doi: 10.1038/nrn1533. PMID: 15496861.

2. Golinkoff RM, Can DD, Soderstrom M, Hirsh-Pasek K. (Baby) Talk to me: The social context of infant-directed speech and its effects on early language acquisition. Curr Dir Psychol Sci. 2015;24(5):339–344. doi: 10.1177/0963721415595345.

3. Holzrichter AS, Meier RP. Child-directed signing in American sign language. In: Chamberlain C, C JPM, Mayberry RI, editors. Language acquisition by eye. Hillsdale, NJ: Lawrence Erlbaum Associates; 2000. p. 25–40.

4. Masataka N. Motherese in a signed language. Infant Behav Dev. 1992;15(4):453–460. doi: 10.1016/0163-6383(92)80013-K.

5. De Boer B. Infant-directed speech and the evolution of language. In: Tallerman M, editor. Evolutionary prerequisites for language. Oxford: Oxford University Press; 2005. p. 100–121.

6. Chong S, Werker JF, Russell JA, Carroll JM. Three facial expressions mothers direct to their infants. Infant Child Dev. 2003;12(3):211–232. doi: 10.1002/icd.286.

7. Iverson JM, Capirci O, Longobardi E, Caselli MC. Gesturing in mother-child interactions. Cogn Dev. 1999;14(1):57–75. doi: 10.1016/S0885-2014(99)80018-5.

8. Brand RJ, Baldwin DA, Ashburn LA. Evidence for 'motionese': modifications in mothers' infant-directed action. Dev Sci. 2002;5(1):72–83. doi: 10.1111/1467-7687.00211.

9. Soderstrom M. Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. Dev Rev. 2007;27(4):501–532. doi: 10.1016/j.dr.2007.06.002.

10. Csibra G, Gergely G. Natural pedagogy. Trends Cogn Sci. 2009;13(4):148–153. doi: 10.1016/j.tics.2009.01.005.

11. Tomasello M. Constructing a language: A usage-based theory of language acquisition. Cambridge, MA: Harvard University Press; 2003.

12. Tomasello M, Striano T, Rochat P. Do young children use objects as symbols? Br J Dev Psychol. 1999;17(4):563–584. doi.org/10.1348/026151099165483.

13. Janik VM, Slater PJ. The different roles of social learning in vocal communication. Anim Behav. 2000;60(1):1-11. doi: 10.1006/anbe.2000.1410. PMID: 10924198.

14. Vernes SC, Kriengwatana BP, Beeck VC, Fischer J, Tyack PL, Ten Cate C, Janik VM. The multi-dimensional nature of vocal learning. Philos Trans R Soc Lond B Biol Sci. 25;376(1836):20200236. doi: 10.1098/rstb.2020.0236. PMID: 34482723.

15. Janik VM, Slater PJ. Vocal learning in mammals. Adv Study Behav. 1997;26(1):59–100.

16. Shneidman LA, LA, Goldin-Meadow S. Language input and acquisition in a Mayan village: How important is directed speech? Dev Sci. 2012;15(5):659–673. doi: 10.1111/j.1467-7687.2012.01168.x. PMID: 22925514.

17. Cristia A, A, Gurven M, Stieglitz J. Child-directed speech is infrequent in a forager-farmer population: A time allocation study. Child Dev. 2019;90(3):759–773. doi: 10.1111/cdev.12974. PMID: 29094348.

18. Snowdon CT. Language capacities of nonhuman animals. Am J Phys Anthropol. 1990;33(S11):215–243. doi: 10.1002/ajpa.1330330510.

19. Talmage-Riggs G, Winter P, Ploog D, Mayer W. Effect of deafening on the vocal behavior of the squirrel monkey (*Saimiri sciureus*). Folia Primatol. 1972;17(5-6):404–420. doi: 10.1159/000155458. PMID: 4628332.

20. Winter P, Handley P, Ploog D, Schott D. Ontogeny of squirrel monkey calls under normal conditions and under acoustic isolation. Vocal development in squirrel monkeys. Behaviour. 1973;47(3-4):230–239. doi: 10.1163/156853973x00085. PMID: 4203637.

21. Hammerschmidt K, Freudenstein T, Jürgens U. Vocal development in squirrel monkeys. Behaviour. 2001;138(9):1179–1204.

22. Seyfarth RM, Cheney DL. Vocal development in vervet monkeys. Anim Behav. 1986;34(6):1640–1658.

23. Zuberbühler K. Interspecies semantic communication in two forest primates. Proc Biol Sci. 2000;267(1444):713-8. doi: 10.1098/rspb.2000.1061. PMID: 10821618.

24. Seyfarth RM, Cheney DL. Production, usage, and comprehension in animal vocalizations. Brain Lang. 2010;115(1):92–100. doi: 10.1016/j.bandl.2009.10.003. PMID: 19944456.

25. Wegdell F, Hammerschmidt K, Fischer J. Conserved alarm calls but rapid auditory learning in monkey responses to novel flying objects. Nat Ecol Evol. 2019;3(7):1039–1042. doi: 10.1038/s41559-019-0903-5. PMID: 31133723.

26. Zimmermann E. The vocal repertoire of the adult Senegal bushbaby (*Galago senegalensis senegalensis*). Behaviour. 1985;94(3/4):212–233.

27. Fischer J, Hammerschmidt K. Towards a new taxonomy of primate vocal production learning. Philos Trans R Soc Lond B Biol Sci. 2020;375(1789):20190045. doi: 10.1098/rstb.2019.0045. PMID: 31735147.

28. Takahashi DY, Fenley AR, Teramoto Y, Narayanan DZ, Borjon JI, Holmes P, et al. The developmental dynamics of marmoset monkey vocal production. Science. 2015;349(6249):734–738. doi: 10.1126/science.aab1058. PMID: 26273055.

29. Watson SK, Townsend SW, Schel AM, Wilke C, Wallace EK, Cheng L, et al. Vocal learning in the functionally referential food grunts of chimpanzees. Curr Biol. 2015;25(4):495-9. doi: 10.1016/j.cub.2014.12.032. PMID: 25660548.

30. Fischer J, Wheeler BC, Higham JP. Is there any evidence for vocal learning in chimpanzee food calls? Curr Biol. 2015;25(21):R1028-R1029. doi: 10.1016/j.cub.2015.09.010. PMID: 26528740.

31. Katsu N, Yamada K, Nakamichi M. Development in the usage and comprehension of greeting calls in a free-ranging group of Japanese macaques (*Macaca fuscata*). Ethology. 2014;120(10):1024–1034. doi: 10.1111/eth.12275.

32. Lameira AR, Hardus ME, Mielke A, Wich SA, Shumaker RW. Vocal fold control beyond the species-specific repertoire in an orang-utan. Sci Rep. 2016;6:30315. doi: 10.1038/srep30315. PMID: 27461756.

33. Laporte MNC, Zuberbühler K. The development of a greeting signal in wild chimpanzees. Dev Sci. 2011;14(5):1220–1234. doi: 10.1111/j.1467-7687.2011.01069.x. PMID: 21884337.

34. Oller DK, Griebel U, Iyer SN, Jhang Y, Warlaumont AS, Dale R, et al. Language origins viewed in spontaneous and interactive vocal rates of human and bonobo infants. Front Psychol. 2019;10:729. doi: 10.3389/fpsyg.2019.00729. PMID: 31001176.

35. Fröhlich M, Wittig RM, Pika S. Play-solicitation gestures in chimpanzees in the wild: flexible adjustment to social circumstances and individual matrices. R Soc Open Sci. 2016;3(8):160278. doi: 10.1098/rsos.160278. PMID: 27853603.

36. Wich SA, Krützen M, Lameira AR, Nater A, Arora N, Bastian ML, et al. Call cultures in orang-utans? PLoS One. 2012;7(5):e36180. doi: 10.1371/journal.pone.0036180. PMID: 22586464.

37. Knox A, Markx J, How E, Azis A, Hobaiter C, van Veen FJF, et al. Gesture use in communication between mothers and offspring in wild orang-utans (*Pongo pygmaeus wurmbii*) from the Sabangau Peat-Swamp Forest, Borneo. Int J of Primatol. 2019;40(3):393–416. doi.org/10.1007/s10764-019-00095-w.

38. Fröhlich M, Wittig RM, Pika S. Should I stay or should I go? Initiation of joint travel in mother–infant dyads of two chimpanzee communities in the wild. Anim Cogn. 2016;19(3):483–500. doi: 10.1007/s10071-015-0948-z. PMID: 26833496.

39. Halina M, Rossano F, Tomasello M. The ontogenetic ritualization of bonobo gestures. Anim Cogn. 2013;16(4):653-66. doi: 10.1007/s10071-013-0601-7. PMID: 23370783.

40. Luef EM, Liebal K. Infant-Directed Communication in Lowland Gorillas (*Gorilla gorilla*): Do older animals scaffold communicative competence in infants? Am J Primatol. 2012;74(9):841-52. doi: 10.1002/ajp.22039. PMID: 22644596.

41. Fröhlich M, Müller G, Zeiträg C, Wittig RM, Pika S. Gestural development of chimpanzees in the wild: the impact of interactional experience. Anim Behav. 2017;134:271–282. doi: 10.1016/j.anbehav.2016.12.018.

42. Fröhlich M, van Schaik C. Social tolerance and interactional opportunities as drivers of gestural redoings in orang-utans. Philos Trans R Soc B. Forthcoming 2022.

43. Genty E, Neumann C, Zuberbühler K. Bonobos modify communication signals according to recipient familiarity. Sci Rep. 2015;5:16442. doi: 10.1038/srep16442. PMID: 26552655.

44. Tomasello M, Call J, Warren J, Frost GT, Carpenter M, Nagell K. The ontogeny of chimpanzee gestural signals: A comparison across groups and generations. EOC. 1997;1(2):223–259. doi: 10.1075/eoc.1.2.04tom.

45. Ferguson CA. Talking to children: A search for universals. In: Greenberg JH, editor. Universals of human language. Standford: Stanford University Press; 1978. p. 203–224.

46. Fernald A, Taeschner T, Dunn J, Papoušek M, De Boysson-Bardies B, Fukui I. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. J Child Lang. 1989 ;16(3):477-501. doi: 10.1017/s0305000900010679. PMID: 2808569.

47. Singh L, Nestor S, Parikh C, Yull A. Influences of infant-directed speech on early word recognition. Infancy. 2009;14(6):654–666. doi: 10.1080/15250000903263973. PMID: 32693515.

48. Song JY, Demuth K, Morgan J. Effects of the acoustic properties of infant-directed speech on infant word recognition. J Acoust Soc Am. 2010;128(1):389-400. doi: 10.1121/1.3419786. PMID: 20649233.

49. Porritt LL, Zinser MC, Bachorowski JA, Kaplan PS. Depression diagnoses and fundamental frequency-based acoustic cues in maternal infant-directed speech. Lang Learn Dev. 2014;10(1):51-67. doi: 10.1080/15475441.2013.802962. PMID: 24489521.

50. Rowe ML. Longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. Child Dev. 2012;83(5):1762-74. doi: 10.1111/j.1467-8624.2012.01805.x. PMID: 22716950.

51. Hartman KM, Ratner NB, Newman RS. Infant-directed speech (IDS) vowel clarity and child language outcomes. J Child Lang. 2017;44(5):1140-1162. doi: 10.1017/S0305000916000520. PMID: 27978860.

52. Tomasello M, Farrar MJ. Joint attention and early language. Child Dev. 1986;57(6):1454-63. PMID: 3802971.

53. Burkart JM, Hrdy SB, Van Schaik CP. Cooperative breeding and human cognitive evolution. Evol Anthropol. 2009;18(5):175–186. doi: 10.1002/evan.20222.

54. Fernald A, Simon T. Expanded intonation contours in mothers' speech to newborns. Dev Psychol. 1984;20(1):104–113. doi: 10.1037/0012-1649.20.1.104.

55. Ratner NB. Durational cues which mark clause boundaries in mother–child speech. J Phon. 1986;14(2):303–309. doi: 10.1016/S0095-4470(19)30670-9.

56. Uther M, Knoll MA, Burnham D. Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. Speech Commun. 2007;49(1):2–7. doi: 10.1016/j.specom.2006.10.003.

57. Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskina VL, et al. Cross-language analysis of phonetic units in language addressed to infants. Science. 1997;277(5326):684–686. doi: 10.1126/science.277.5326.684. PMID: 9235890.

58. Fernald A. Intonation and communicative intent in mothers' speech to infants: Is the melody the message? Child Dev. 1989;60(6):1497–1510. PMID: 2612255.

59. Englund KT. Voice onset time in infant directed speech over the first six months. First Lang. 2005;25(2):219–234. doi: 10.1177/0142723705050286.

60. Cooper RP, Aslin RN. Preference for infant-directed speech in the first month after birth. Child Dev. 1990;61(5):1584–1595. PMID: 2245748.

61. Papoušek M, Papoušek H, Haekel M. Didactic adjustments in fathers' and mothers' speech to their 3-month-old infants. J Psycholinguist Res. 1987;16(5):491–516.

62. Cameron-Faulkner T, Lieven E, Tomasello M. A construction based analysis of child directed speech. Cogn Sci. 2003;27(6):843–873. doi: 10.1207/s15516709cog2706_2.

63. Snow CE. The development of conversation between mothers and babies. J Child Lang. 1977;4(1):1–22. doi:10.1017/S0305000900000453.

64. Martin A, Igarashi Y, Jincho N, Mazuka R. Utterances in infant-directed speech are shorter, not slower. Cognition. 2016;156:52–59. doi: 10.1016/j.cognition.2016.07.015.

65. Phillips JR. Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. Child Dev. 1973;44(1):182–185.

66. Henninga A, Striano T, Lieven EVM. Maternal speech to infants at 1 and 3 months of age. Infant Behav Dev. 2005;28(4):519–536. doi: 10.1016/j.infbeh.2005.06.001.

67. Kempe V, Brooks PJ, Pirott L. How can child-directed speech facilitate the acquisition of morphology? In: Research on child language acquisition: Proceedings of the 8th conference of the International Association for the Study of Child Language. Cascadilla Press; 2001. p. 1237–1247.

68. Kempe V, Brooks PJ, Gillis S. Diminutives provide multiple benefits for language acquisition. In: Savickienė I, Dressler WU, editors. The Acquisition of diminutives: A cross-linguistic perspective. Amsterdam: John Benjamins B. V.; 2007. p. 319–342.

69. Newport E, Gleitman H, Gleitman L. Mother, I'd rather do it myself: Some effects and non-effects of motherese. In: Snow CE, Ferguson CA, editors. Talking to children. Cambridge: Cambridge University Press; 1977. p. 109–149.

70. Soderstrom M, Blossom M, Foygel R, Morgan JL. Acoustical cues and grammatical units in speech to two preverbal infants. J Child Lang. 2008;35(4):869-902. doi: 10.1017/S0305000908008763. PMID: 18838016.

71. Küntay A, Slobin DI. Listening to a Turkish Mother: Some Puzzles for Acquisition. In: Slobin DI, Gerhardt J, Kyratzis A, Guo J, editors. Social interaction, social context, and language: essays in honor of Susan-Ervin Tripp. Hillsdale, NJ: Lawrence Erlbaum Associates; 1996. p. 265–286.

72. Lester NA, Moran S, Küntay AC, Allen SEM, Pfeiler B, Stoll S. Detecting structured repetition in child-surrounding speech: Evidence from maximally diverse languages. Cognition. 2022;221:104986. doi: 10.1016/j.cognition.2021.104986. PMID: 34953269.

73. Yurovsky D, Yu C, Smith LB. Statistical speech segmentation and word learning in parallel: scaffolding from child-directed speech. Front Psychol. 2012;3:374. doi: 10.3389/fpsyg.2012.00374. PMID: 23162487.

74. Albin DD, Echols CH. Stressed and word-final syllables in infant-directed speech. Infant Behav Dev. 1996;19(4):401–418. doi.org/10.1016/S0163-6383(96)90002-8.

75. Trainor LJ, Desjardins RN. Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. Psychon Bull Rev. 2002;9(2):335-40. doi: 10.3758/bf03196290. PMID: 12120797.

76. Werker JF, McLeod PJ. Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. Can J Psychol. 1989;43(2):230-46. doi: 10.1037/h0084224. PMID: 2486497.

77. Cohen D, Cassel RS, Saint-Georges C, Mahdhaoui A, Laznik MC, Apicella F, et al. Do parentese prosody and fathers' involvement in interacting facilitate social interaction in infants who later develop autism? PLoS One. 2013;8(5):e61402. doi: 10.1371/journal.pone.0061402. PMID: 23650498.

78. Dunst CJ, Gorman E, Hamby DW. Preference for infant-directed speech in preverbal young children. CELL. 2012;5(1).

79. Fernald A. Four-month-old infants prefer to listen to motherese. Infant Behav Dev. 1985;8(2):181–195. doi: 10.1016/S0163-6383(85)80005-9.

80. Naoi N, Minagawa-Kawai Y, Kobayashi A, Takeuchi K, Nakamura K, Yamamoto JI, et al. Cerebral responses to infant-directed speech and the effect of talker familiarity. NeuroImage. 2012;59(2):1735–1744. doi: 10.1016/j.neuroimage.2011.07.093. PMID: 21867764.

81. Niwano K, Sugai K. Acoustic determinants eliciting Japanese infants' vocal response to maternal speech. Psychol Rep. 2002;90(1):83-90. doi: 10.2466/pr0.2002.90.1.83. PMID: 11899017.

82. Karzon RG. Discrimination of polysyllabic sequences by one- to four-month-old infants. J Exp Child Psychol. 1985;39(2):326-42. doi: 10.1016/0022-0965(85)90044-x. PMID: 3989467.

83. Bortfeld H, Morgan JL. Is early word-form processing stress-full? How natural variability supports recognition. Cogn Psychol. 2010;60(4):241-66. doi: 10.1016/j.cogpsych.2010.01.002. PMID: 20159653.

84. Lyakso EE, Frolova OV, Grigorev AS. Infant vocalizations at the first year of life predict speech development at 2-7 years: Longitudinal study. Psychology. 2014;5(12). doi: 10.4236/psych.2014.512154.

85. Thiessen ED, Hill EA, Saffran JR. Infant-directed speech facilitates word segmentation. Infancy. 2005;7(1):53–71. doi: 10.1207/s15327078in0701_5. PMID: 33430544.

86. Herold DS, Nygaard LC, Namy LL. Say it like you mean it: Mothers' use of prosody to convey word meaning. Lang Speech. 2012;55(Pt 3):423-36. doi: 10.1177/0023830911422212. PMID: 23094322.

87. Spinelli M, Fasolo M, Mesman J. Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. Dev Rev. 2017;44:1–18. doi.org/10.1016/j.dr.2016.12.001.

88. Ambridge B, Kidd E, Rowland CF, Theakston AL. The ubiquity of frequency effects in first language acquisition. J Child Lang. 2015;42(2):239-73. doi: 10.1017/S030500091400049X. PMID: 25644408.

89. Lieven E. Input and first language acquisition: Evaluating the role of frequency. Lingua. 2010;120(11):2546–2556. doi: 10.1016/j.lingua.2010.06.005.

90. Stoll S, Abbot-Smith K, Lieven E. Lexically Restricted Utterances in Russian, German, and English child-directed speech. Cogn Sci. 2009;33(1):75-103. doi: 10.1111/j.1551-6709.2008.01004.x. PMID: 21585464.

91. Mintz TH. Frequent frames as a cue for grammatical categories in child directed speech. Cognition. 2003;90:91–117. doi: 10.1016/s0010-0277(03)00140-9. PMID: 14597271.

92. Moran S, Blasi DE, Schikowski R, Küntay AC, Pfeiler B, Allen S, et al. A universal cue for grammatical categories in the input to children: Frequent frames. Cognition. 2018;175:131–140. doi: 10.1016/j.cognition.2018.02.005. PMID: 29518682.

93. Hills T. The company that words keep: comparing the statistical structure of child-versus adult-directed language. J Child Lang. 2013;40(3):586-604. doi: 10.1017/S0305000912000165. PMID: 22584041.

94. You G, Bickel B, Daum MM, Stoll S. Child-directed speech is optimized for syntax-free semantic inference. Sci Rep. 2021;11(1):16527. doi: 10.1038/s41598-021-95392-x. PMID: 34400656.

95. Fernald A, Hurtado N. Names in frames: Infants interpret words in sentence frames faster than words in isolation. Dev Sci. 2006;9(3):33–40. doi: 10.1111/j.1467-7687.2006.00482.x. PMID: 16669790.

96. Lew-Williams C, Pelucchi B, Saffran JR. Isolated words enhance statistical language learning in infancy. Dev Sci. 2011;14(6):1323–1329. doi: 10.1111/j.1467-7687.2011.01079.x. PMID: 22010892.

97. Waterfall HR. A little change is a good thing: Feature theory, language acquisition and variation sets [dissertation]. Chicago: University of Chicago; 2006.

98. Schwab JF, Lew-Williams C. Language learning, socioeconomic status, and child-directed speech. Wiley Interdiscip Rev Cogn Sci. 2016;7(4):264-75. doi: 10.1002/wcs.1393. PMID: 27196418.

99. Hart B, Risley TR. Meaningful differences in the everyday experience of young American children. Baltimore: Paul H Brookes Publishing; 1995.

100. Pan BA, Rowe ML, Singer JD, Snow CE. Maternal correlates of growth in toddler vocabulary production in low-income families. Child Dev. 2005;76(4):763-82. doi: 10.1111/j.1467-8624.2005.00876.x. PMID: 16026495.

101. Weisleder A, Fernald A. Talking to children matters: Early language experience strengthens processing and builds vocabulary. Psychol Sci. 2013;24(11):2143-52. doi: 10.1177/0956797613488145. PMID: 24022649.

102. Newman RS, Rowe ML, Bernstein Ratner N. Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. J Child Lang. 2016;43(5):1158-73. doi: 10.1017/S0305000915000446. PMID: 26300377.

103. Huttenlocher J, Waterfall H, Vasilyeva M, Vevea J, Hedges LV. Sources of variability in children's language growth. Cogn Psychol. 2010;61(4):343-65. doi: 10.1016/j.cogpsych.2010.08.002. PMID: 20832781.

104. Hirsh-Pasek K, Adamson LB, Bakeman R, Owen MT, Golinkoff RM, Pace A, et al. The contribution of early communication quality to low income children's language success. Psychol Sci. 2015;26(7):1071-83. doi: 10.1177/0956797615581493. PMID: 26048887.

105. Lieven EV. Crosslinguistic and crosscultural aspects of language addressed to children. In: Gallaway C, Richards BJ, editors. Input and interaction in language acquisition. Cambridge: Cambridge University Press; 1994. p. 56–73.

106. Ochs E, Schieffelin, Bambi B. Language acquisition and socializations: Three developmental stories and their implications. In: Shweder RA, LeVine RA, editors. Culture theory: Essays on mind, self and emotion. Cambridge: Cambridge University Press; 1984. p. 276–320.

107. Henrich J, Heine SJ, Norenzayan A. Most people are not WEIRD. Nature. 2010;466(7302):29–29. doi: 10.1038/466029a.

108. Vogt P, Mastin JD, Schots DM. Communicative intentions of child-directed speech in three different learning environments: Observations from the Netherlands, and rural and urban Mozambique. First Lang. 2015;35(4-5):341–358. doi: 10.1177/0142723715596647.

109. Casillas M, Brown P, Levinson SC. Early language experience in a Tseltal Mayan village. Child Dev. 2020;91(5):1819-1835. doi: 10.1111/cdev.13349. PMID: 31891183.

110. Bergelson E, Casillas M, Soderstrom M, Seidl A, Warlaumont AS, Amatuni A. What do North American babies hear? A large-scale cross-corpus analysis. Dev Sci. 2019;22(1):e12724. doi: 10.1111/desc.12724. PMID: 30369005.

111. Keller H. Cultures of infancy. Mahwah, NJ: Lawrence Erlbaum Associates; 2013.

112. Sarvasy H, Elvin J, Li W, Escudero P. An acoustic analysis of Nungon vowels in child-versus adult-directed speech. In: Proceedings of the 19th International Congress of Phonetic Sciences Melbourne; 2019. p. 3155–3159.

113. Broesch TL, Bryant GA. Prosody in infant-directed speech Is similar across Western and traditional cultures. J Cogn Dev. 2015;16(1):31–43. doi.org/10.1080/15248372.2013.833923.

114. Pye C. Quiché Mayan speech to children. J Child Lang. 1986;13(1):85-100. doi: 10.1017/s0305000900000313. PMID: 3949901.

115. Farran LK, Lee CC, Yoo H, Oller DK. Cross-cultural register differences in infant-directed speech: An initial study. PLoS One. 2016;11(3):e0151518. doi: 10.1371/journal.pone.0151518. PMID: 26981626.

116. Ratner NB, Pye C. Higher pitch in BT is not universal: Acoustic evidence from Quiché Mayan. J Child Lang. 1984;11(3):515-22. doi: 10.1017/s0305000900005924. PMID: 6501462.

117. Stern DN, Spieker S, Barnett RK, MacKain K. The prosody of maternal speech: Infant age and context related changes. J Child Lang. 1983;10(1):1-15. doi: 10.1017/s0305000900005092. PMID: 6841483.

118. Vosoughi S, Roy D. A longitudinal study of prosodic exaggeration in child-directed speech in Speech Prosody. In: 6th International Conference. SProSIG; 2012.

119. Kondaurova MV, Bergeson TR, Xu H. Age-related changes in prosodic features of maternal speech to prelingually deaf infants with cochlear implants. Infancy. 2013;18(5):825–848. doi: 10.1111/infa.12010. PMID: 24244108.

120. Narayan CR, McDermott LC. Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. J Acoust Soc Am. 2016;139(3):1272-81. doi: 10.1121/1.4944634. PMID: 27036263.

121. Fernald A. Meaningful melodies in mothers’ speech to infants. In: Papoušek H, Jürgens U, Papoušek M, editors. Nonverbal vocal behaviour. Cambridge: Cambridge University Press; 1992. p. 262–282.

122. Gratier M, Devouche E, Guellai B, Infanti R, Yilmaz E, Parlato-Oliveira E. Early development of turn-taking in vocal interaction between mothers and infants. Front Psychol. 2015;6:1167. doi: 10.3389/fpsyg.2015.01167. PMID: 26388790.

123. Zimmermann E. Aspects of reproduction and behavioral and vocal development in Senegal bushbabies (*Galago senegalensis*). Int J Primatol. 1989;10(1):1–16. doi.org/10.1007/BF02735700.

124. Biben M, Symmes D, Bernhards D. Contour variables in vocal communication between squirrel monkey mothers and infants. Dev Psychobiol. 1989;22(6):617-31. doi: 10.1002/dev.420220607. PMID: 2792572.

125. Szenczi P, Bánszegi O, Urrutia A, Faragó T, Hudson R. Mother-offspring recognition in the domestic cat: Kittens recognize their own mother's call. Dev Psychobiol. 2016;58(5):568-77. doi: 10.1002/dev.21402. PMID: 26935009.

126. Sèbe F, Duboscq J, Aubin T, Ligout S, Poindron P. Early vocal recognition of mother by lambs: contribution of low- and high-frequency vocalizations. Anim Behav. 2010;79(5):1055–1066. doi: 10.1016/j.anbehav.2010.01.021.

127. Padilla De La Torre M, Briefer EF, Ochocki BM, McElligott AG, Reader T. Mother-offspring recognition via contact calls in cattle, Bos taurus. Anim Behav. 2016;114:147–154. doi: 10.1016/j.anbehav.2016.02.004.

128. Volodin IA, Sibiryakova OV, Soldatova NV, Volodina EV. Acoustically different contact calls of mother and young goitred gazelle are equally individualized: Is this a common relationship in ruminants? Behaviour. 2019;156(12):1185–1207. doi: 10.1163/1568539X-00003561.

129. Sibiryakova OV, Volodin IA, Frey R, Zuther S, Kisebaev TB, Salemgareev AR, et al. Remarkable vocal identity in wild-living mother and neonate saiga antelopes: A specialization for breeding in huge aggregations? Sci Nat. 2017;104(3-4):11. doi: 10.1007/s00114-017-1433-0. PMID: 28243711.

130. Balcombe JP, McCracken GF. Vocal recognition in Mexican free-tailed bats: do pups recognize mothers? Anim Behav. 1992;43(1):79–87. doi: 10.1016/S0003-3472(05)80961-3.

131. Charrier I, Mathevon N, Jouventin P. Mother’s voice recognition by seal pups. Nature. 2001;412(6850):873. doi: 10.1038/35091136. PMID: 11528465.

132. Ruch H, Zürcher Y, Burkart JM. The function and mechanism of vocal accommodation in humans and other primates. Biol Rev Camb Philos Soc. 2018;93(2):996-1013. doi: 10.1111/brv.12382. PMID: 29111610.

133. Weiss BM, Ladich F, Spong P, Symonds H. Vocal behavior of resident killer whale matrilines with newborn calves: The role of family signatures. J Acoust Soc Am. 2006;119(1):627-35. doi: 10.1121/1.2130934. PMID: 16454316.

134. Zürcher Y, Willems EP, Burkart JM. Trade-offs between vocal accommodation and individual recognisability in common marmoset vocalizations. Sci Rep. 2021;11(1):15683. doi: 10.1038/s41598-021-95101-8. PMID: 34344939.

135. Castro NA, Snowdon CT. Development of vocal responses in infant cotton-top tamarins. Behaviour. 2000;137(5):629–646. doi: 10.1163/156853900502259.

136. Koda H, Lemasson A, Oyakawa C, Pamungkas J, Masataka N. Possible role of mother-daughter vocal interactions on the development of species-specific song in gibbons. PLoS One. 2013;8(8):e71432. doi: 10.1371/journal.pone.0071432. PMID: 23951160.

137. Takahashi DY, Liao DA, Ghazanfar AA. Vocal learning via social reinforcement by infant marmoset monkeys. Curr Biol. 2017;27(12):1844-1852.e6. doi: 10.1016/j.cub.2017.05.004. PMID: 28552359.

138. Chow CP, Mitchell JF, Miller CT. Vocal turn-taking in a non-human primate is learned during ontogeny. Proc Biol Sci B. 2015;282(1807):20150069. doi: 10.1098/rspb.2015.0069. PMID: 25904663.

139. Takahashi DY, Fenley AR, Ghazanfar AA. Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. Philos Trans R Soc Lond B Biol Sci. 2016;371(1693):20150370. doi: 10.1098/rstb.2015.0370. PMID: 27069047.

140. Chen Y, Matheson LE, Sakata JT. Mechanisms underlying the social enhancement of vocal learning in songbirds. Proc Natl Acad Sci USA. 2016;113(24):6641-6. doi: 10.1073/pnas.1522306113. PMID: 27247385.

141. Fernandez AA, Knörnschild M. Pup directed vocalizations of adult females and males in a vocal learning bat. Front Ecol Evol. 2020;8:265. doi: 10.3389/fevo.2020.00265.

142. Koda H. Possible use of heterospecific food-associated calls of macaques by sika deer for foraging efficiency. Behav Processes. 2012;91(1):30-4. doi: 10.1016/j.beproc.2012.05.006. PMID: 22641112.

143. Brown P. The Cultural organization of attention. In: Duranti A, Ochs E, Schieffelin BB, editors. The handbook of language socialization. Amsterdam: John Benjamins B.V.; 2011. p. 29–55.

144. De León L. Language socialization and multiparty participation frameworks. In: A Duranti EO, Schieffelin BB, editors. The handbook of language socialization. Amsterdam: John Benjamins B.V.; 2011. p. 81–111.

145. Ivey PK. Cooperative reproduction in Ituri forest hunter-gatherers: Who cares for Efe infants? Curr Anthropol. 2000;41(5):856–866. doi: 10.1086/317414

146. Casillas M, Brown P, Levinson SC. Early language experience in a Papuan community. J Child Lang. 2021;48(4):792-814. doi: 10.1017/S0305000920000549. PMID: 32988426.

147. Hobaiter C, Byrne RW, Zuberbühler K. Wild chimpanzees' use of single and combined vocal and gestural signals. Behav Ecol Sociobiol. 2017;71(6):96. doi: 10.1007/s00265-017-2325-1. PMID: 28596637.

148. Carpenter M, Nagell K, Tomasello M. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monogr Soc Res Child Dev. 1998;63(4):i-vi, 1-143.

149. Scott K, Sakkalou E, Ellis-Davies K, Hilbrink E, Hahn U, Gattis M. Infant contributions to joint attention predict vocabulary development. In: CogSci 2013: the 35th Annual Conference of the Cognitive Science Society. Cognitive Science Society; 2013. p. 3384–3389.

150. Rowland CF, Pine JM, Lieven EVM, Theakston AL. Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech. J Child Lang. 2003;30(3):609-35. PMID: 14513470.