Manuscript version: Author’s Accepted Manuscript
The version presented in WRAP is the author’s accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:
http://wrap.warwick.ac.uk/165271

How to cite:
Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher’s statement:
Please refer to the repository item page, publisher’s statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.
The function and evolution of child-directed communication

Johanna Schick\textsuperscript{1,5¶}, Caroline Fryns\textsuperscript{2,5¶}, Franziska Wegdel\textsuperscript{1,5¶}, Marion Laporte\textsuperscript{7,8}, Klaus Zuberbühler\textsuperscript{2,4,5}, Carel P. van Schaik\textsuperscript{3,5,9&*}, Simon W. Townsend\textsuperscript{1,5,6&}, Sabine Stoll\textsuperscript{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.
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,
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).

<table>
<thead>
<tr>
<th>Type of feature</th>
<th>Known feature of CDC</th>
<th>Proposed function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acoustic</td>
<td>Pitch variability</td>
<td>Attention grabbing</td>
<td>[54]</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Lengthening of vowels and pauses</td>
<td>Segmentation and discrimination of sounds</td>
<td>[55,56]</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Extended vowel triangle</td>
<td>Sound discrimination</td>
<td>[57]</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Clear articulation</td>
<td>Facilitate comprehension</td>
<td>[46,58]</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Increased voice-onset time</td>
<td>Sound discrimination</td>
<td>[59]</td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
<td>Benefits</td>
<td>References</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Acoustic</td>
<td>Slower speaking rate</td>
<td>Facilitate comprehension, discrimination and segmentation</td>
<td>[54,60]</td>
</tr>
<tr>
<td>Structural</td>
<td>Frequent repetitions</td>
<td>Structural generalization of word/unit classes</td>
<td>[61,62]</td>
</tr>
<tr>
<td>Structural</td>
<td>Short utterances</td>
<td>Facilitate comprehension</td>
<td>[63,64]</td>
</tr>
<tr>
<td>Structural</td>
<td>Low type/token ratio</td>
<td>Facilitate comprehension</td>
<td>[65,66]</td>
</tr>
<tr>
<td>Structural</td>
<td>Simplified syntax and semantics</td>
<td>Facilitate comprehension</td>
<td>[63,65]</td>
</tr>
<tr>
<td>Structural</td>
<td>Frequent use of diminutives</td>
<td>Simplification of certain morphological aspects (language specific)</td>
<td>[67,68]</td>
</tr>
<tr>
<td>Structural</td>
<td>Frequent questions</td>
<td>Invite response, repetition, attention grabbing</td>
<td>[69,70]</td>
</tr>
<tr>
<td>Structural</td>
<td>Variation sets</td>
<td>Structural generalization of word classes</td>
<td>[71,72]</td>
</tr>
<tr>
<td>Structural</td>
<td>Scaffolding</td>
<td>Learning of word constructions</td>
<td>[73]</td>
</tr>
</tbody>
</table>

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 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


dynamics of marmoset monkey vocal production. Science. 2015;349(6249):734–738. doi:
10.1126/science.aab1058. PMID: 26273055.


30. Fischer J, Wheeler BC, Higham JP. Is there any evidence for vocal learning in chimpanzee food calls?

31. Katsu N, Yamada K, Nakamichi M. Development in the usage and comprehension of greeting calls in

32. Lameira AR, Hardus ME, Mielke A, Wich SA, Shumaker RW. Vocal fold control beyond the species-
27461756.


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.


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.


of prosodic modifications in mothers’ and fathers’ speech to preverbal infants. J Child Lang. 1989

47. Singh L, Nestor S, Parikh C, Yull A. Influences of infant-directed speech on early word recognition.

48. Song JY, Demuth K, Morgan J. Effects of the acoustic properties of infant-directed speech on infant

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:

50. Rowe ML. Longitudinal investigation of the role of quantity and quality of child-directed speech in
PMID: 22716950.


3802971.


doi:10.1017/S0305000900000453.

64. Martin A, Igarashi Y, Jincho N, Mazuka R. Utterances in infant-directed speech are shorter, not


66. Henninga A, Striano T, Lieven EVM. Maternal speech to infants at 1 and 3 months of age. Infant

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

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.

69. Newport E, Gleitman H, Gleitman L. Mother, I’d rather do it myself: Some effects and non-effects of

70. Soderstrom M, Blossom M, Foygel R, Morgan JL. Acoustical cues and grammatical units in speech to
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


81. Niwano K, Sugai K. Acoustic determinants eliciting Japanese infants’ vocal response to maternal

82. Karzon RG. Discrimination of polysyllabic sequences by one- to four-month-old infants. J Exp Child

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.

86. Herold DS, Nygaard LC, Namy LL. Say it like you mean it: Mothers’ use of prosody to convey word

87. Spinelli M, Fasolo M, Mesman J. Does prosody make the difference? A meta-analysis on relations

88. Ambridge B, Kidd E, Rowland CF, Theakston AL. The ubiquity of frequency effects in first language


124. Biben M, Symmes D, Bernhards D. Contour variables in vocal communication between squirrel
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

vocal identity in wild-living mother and neonate saiga antelopes: A specialization for breeding in

130. Balcombe JP, McCracken GF. Vocal recognition in Mexican free-tailed bats: do pups recognize

2001;412(6850):873. doi: 10.1038/35091136. PMID: 11528465.


