



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/183087/>

Version: Accepted Version

---

**Article:**

Liebal, Katja, Slocombe, Katie and Waller, Bridget M (2022) The language void 10 years on: multimodal primate communication research is still uncommon. *Ethology ecology*. pp. 274-287. ISSN: 0394-9370

<https://doi.org/10.1080/03949370.2021.2015453>

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

1 **The language void 10 years on: multimodal primate communication research is still**  
2 **uncommon**

3

4 KATJA LIEBAL<sup>1,\*</sup>, KATIE E. SLOCOMBE<sup>2</sup> and BRIDGET M. WALLER<sup>3</sup>

5

6 <sup>1</sup>*Life Sciences, Institute of Biology, Leipzig University, Talstrasse 33, Leipzig 04103, Germany*

7 <sup>2</sup>*Department of Psychology, University of York, York YO10 5DD, UK*

8 <sup>3</sup>*School of Social Sciences, Nottingham Trent University, Shakespeare Street, Nottingham*

9 *NG1 4FQ, UK*

10

11 **Highlights**

- 12 - Primate communication research is still largely unimodal  
13 - Different approaches are used across modalities  
14 - Comparing findings across studies is difficult

15

16 Human language is thought to have evolved from non-linguistic communication systems  
17 present in the primate lineage. Scientists rely on data from extant primate species to  
18 estimate how this happened, with debates centering around which modality (vocalization,  
19 gesture, facial expression) was a likely precursor. In 2011, we demonstrated that different  
20 theoretical and methodological approaches are used to collect data about each modality,  
21 rendering datasets incomplete and comparisons problematic (Slocombe et al. 2011). Here,  
22 10 years later, we conducted a follow-up systematic review to test whether patterns have  
23 changed, examining the primate communication literature published between 2011 and  
24 2020. In sum, despite the promising progress in addressing some gaps in our knowledge,

25 systematic biases still exist and multimodal research remains uncommon. We argue that  
26 theories of language evolution are unlikely to advance until the field of primate  
27 communication research acknowledges and rectifies the gaps in our knowledge.

28

29 \*Corresponding author: Katja Liebal, Human Biology & Primate Cognition, Life Sciences,  
30 Institute of Biology, Leipzig University, Talstrasse 33, 04103 Leipzig, Germany (E-mail:  
31 katja.liebal@uni-leipzig.de).

32

33 KEY WORDS: multimodal, vocalization, facial expression, gesture, primates, language evolution.

34

35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58

## INTRODUCTION

To disentangle the origins and evolution of human language, many researchers investigate our closest relatives, the nonhuman primates (hereafter, primates), with the aim to learn more about which aspects of their communication are shared with humans, and which are unique to humans (e.g., Fitch 2005; Arbib et al. 2008; Tomasello 2008). While speech is clearly limited to the human species, non-verbal means of communication, such as gestural, facial and vocal signals, are shared across a variety of primate species including humans. Therefore, comparative researchers have focused on the communicative behaviors and cognitive skills underlying primate communication to identify potential precursors to human language (Call & Tomasello 2007; Fedurek & Slocombe 2011; Zuberbühler 2015; Seyfarth & Cheney 2017). Common approaches include the analysis of communicative repertoires, the investigation of intentional and referential use, social function, and if/how single signals are combined into longer, possibly meaningful sequences (for an overview, see Liebal et al. 2013).

Comparative researchers interested in the similarities between non-verbal communication of nonhuman primates and humans traditionally focus on either gestures, vocalizations, or facial expressions. These unimodal approaches tend to be associated with fundamentally different theoretical frameworks, which fuels the fierce debates about the most promising precursor of human language (vocal, facial or gestural). Thus, theories supporting a vocal origin of language suggest that language built directly on the vocal abilities of our ancestors, relying on the evidence for referential use of vocalizations and meaningful call combinations in nonhuman primates (Seyfarth 2005; Zuberbühler 2005). Theories proposing a gestural origin suggest that spoken language was preceded by a gestural stage using visual, voluntarily controlled signals (Hewes 1992; Corballis 2002), and

59 highlight the intentional and flexible use of gestures in nonhuman primates (Call &  
60 Tomasello 2007). In contrast to gestures, facial expressions and vocalizations are often  
61 perceived as involuntary expressions of internal affective states; a view supported by the  
62 limited evidence for the learning of novel calls or facial expressions, indicating relatively  
63 closed communicative repertoires (Tomasello 2008). Theories suggesting a facial origin of  
64 language refer to evidence for the speech-like rhythm of communicative mouth movements  
65 in nonhuman primates in support of the hypothesis that such mouth movements represent  
66 precursors to human speech (Bergman 2013; Pereira et al. 2020).

67         However, it is possible that there are systematic differences in how research is  
68 conducted across the different means of communication, rendering strong conclusions and  
69 comparisons about the cognitive features of each problematic. To test this, in 2011, we  
70 conducted a systematic review of primate communication literature covering almost five  
71 decades of research (1960-2008) (Slocombe et al. 2011), with focus on the study of  
72 vocalizations, gestures and facial expressions. As comparative psychologists, we were  
73 interested in the relationship between language evolution and these types of non-verbal  
74 signals, and their potential role in the emergence of human language. As these different  
75 signal types may have different cognitive underpinnings, we used the term “modality” to  
76 refer to vocal, gestural and facial signals. This is different to behavioral ecology approaches,  
77 where modality is defined by the sensory channel through which the signal is received in the  
78 receiver (e.g. visual or auditory channel; Rowe 1999; Partan & Marler 2005; Higham &  
79 Hebets 2013). The debate continues as to how best to label these different types of  
80 communicative signals (Fröhlich et al. 2019), but for consistency with our previous work, we  
81 will refer to vocal, gestural and facial modalities.

82 Slocombe et al.'s (2011) review resulted in two major findings: first, the vast majority  
83 of research studied only one communicative modality (vocalizations, facial expressions, or  
84 gestures), while multimodal approaches investigating two or more modalities and their  
85 interactions in an integrated way were rather the exception than the norm (5%). Second,  
86 facial, gestural and vocal research each relied on rather different theoretical and  
87 methodological approaches. Gestural communication was mainly studied in great apes,  
88 mostly in captive settings using both experimental and observational methods, with a focus  
89 on the producer of a gestural signal. Facial expressions were mostly studied in monkeys, also  
90 mostly in captive settings using observational methods, with a focus on both the producer  
91 and receiver. Vocalizations were also mostly studied in monkeys, in both wild and captive  
92 populations, typically with experimental methods and with a focus on both producer and  
93 receiver. This means that despite the wealth of studies, our review found a lack of facial and  
94 vocal research on apes, gestural research in wild populations, and experimental approaches  
95 to facial communication. Across modalities, there was also a lack of research with a focus on  
96 receivers.

97 These findings had two important implications: first, there were still considerable gaps  
98 in our knowledge about primate communication, and second, findings across modalities  
99 were difficult to compare since different theoretical approaches and methods had been  
100 used. Therefore, we questioned whether the claims regarding a specific origin of human  
101 language – either vocal, gestural, or facial – are legitimate given the existing body of  
102 evidence and the unimodal approaches used to study primate communication (Slocombe et  
103 al. 2011).

104 We also proposed several ways of obtaining a more complete picture of primate  
105 communication and the potential role of the different modalities for the evolution of human

106 language. First, based on the identified gaps of knowledge, we suggested to specifically  
107 target the blind spots, and to conduct more gestural research with monkeys and apes in wild  
108 settings, to conduct more vocal and facial research with great apes, and to focus more on  
109 receiver behavior across modalities. Second, although unimodal research will continue to be  
110 the only option because of methodological constraints, we proposed that "... combining  
111 data, ideas and theories from different modalities might yield a better understanding than  
112 each can provide alone" (Slocombe et al. 2011, p. 920) and therefore suggested a more  
113 integrated, multimodal approach to primate communication, especially where established  
114 methods are available (Liebal et al. 2013). Several recent review papers on multimodal  
115 communication seem to indicate a growing theoretical consensus on the value of  
116 considering a more holistic, multimodal approach to studying communication (Wacewicz &  
117 Zywczyński 2017; Fröhlich & van Schaik 2018; Fröhlich et al. 2019; Singletary & Tecot 2020);  
118 however, whether that has been matched by a growth in empirical studies is not yet known.

119 Therefore, the aim of the current paper was to investigate the current state of the art  
120 in primate communication research 10 years after our publication pointing to these gaps of  
121 knowledge. We conducted a systematic literature review of primate communication  
122 research from 2011-2020 using the same procedure as in Slocombe et al. (2011). We aimed  
123 to examine if (i) the calls for more integrated multimodal work had been answered and  
124 there had been an increase in studies examining two or more modalities and their  
125 interactions from 2011-2020 compared to 1960-2008. Next, we focused on unimodal  
126 research to investigate if the gaps of knowledge had been addressed with (ii) an increase in  
127 gestural research in non-great apes as well as an increase in great ape vocal and facial  
128 research; (iii) an increase in experimental approaches in facial research and observational

129 methods in vocal studies; (iv) an increase of gestural and facial research in wild settings; and  
130 (v) an increase on signal perception and receiver behavior across modalities.

131

## 132 METHODS

133 For our systematic literature search, we searched for literature published between  
134 2011-2020. We used only two of the previously three data bases (Web of Science and  
135 Science Direct), since the third (PrimateLit) has been deaccessioned in 2018. We used the  
136 same search terms as in Slocombe et al. (2011): “facial communication OR facial expression\*  
137 OR facial display OR gestur\* OR gestur\* communication OR gestur\* display OR vocalisation  
138 OR vocalization OR call\* OR vocal communication OR vocal\*” AND “primate\* OR ape\* OR  
139 monkey\* OR macaque\* OR gorilla\* OR baboon\* OR vervet OR chimpanzee OR gibbon\*”.  
140 This resulted in a total of 501 publications.

141 From these search results, we excluded studies that did not examine one or several of  
142 the three modalities of interest (vocal, gestural and facial communication), therefore  
143 excluding studies assessing olfactory or chemical communication only, and any publications  
144 that did not address the topic of primate communication. We also excluded articles that did  
145 not report original empirical research, used secondary data or were not the primary medium  
146 of publication (reviews, meta-analyses, meeting abstracts, conference proceedings, and  
147 book chapters). Computational models and machine learning approaches were only included  
148 if they used original empirical data. We did not consider technical reports (e.g., how to  
149 record vocalizations properly) or biomonitoring (e.g., use of vocalizations to estimate  
150 population size), since they did not specifically target primate communication. Finally, we  
151 excluded studies investigating primates’ responses to human signals or the use of artificial  
152 language systems, while we included studies where primates signaled towards humans.

153           This resulted in a dataset of 294 publications. We coded each of these papers using the  
154 same criteria as in Slocombe et al. (2011). For each publication, we assessed the investigated  
155 “modality” (vocal, gestural, facial or multimodal). Vocal communication included  
156 vocalizations or calls usually produced by the vocal cords and specific sounds produced by  
157 other body parts, such as whistles or raspberries. Gestural communication involved visual  
158 movements of the limbs, head or body postures, but not facial expressions. If they were  
159 manual behaviors accompanied by sounds, such as chest beats, they were also considered  
160 gestures. Facial communication included communicative movements of the face (facial  
161 expressions) or the mouth specifically (sometimes termed orofacial movements or facial  
162 gestures).

163           We coded a study as “multimodal” if it investigated more than one modality, and as  
164 “multimodal integrated” if it investigated the interaction between signals from different  
165 modalities. Some studies investigated facial movements while primates were vocalizing. We  
166 considered them multimodal, but not as instances of integrated multimodal communication,  
167 as these two modalities are necessarily linked with each other via a common production  
168 mechanism. We further examined the “species class” studied (great ape, lesser ape, monkey  
169 or prosimian), whether the “research method” used was observational (no manipulation of  
170 specific variables, no control conditions) or experimental, and whether the “research focus”  
171 was on the producer or receiver of a signal. We also coded the “research environment” and  
172 distinguished between wild (free-ranging individuals in their natural habitats) and captive  
173 settings (laboratories, zoos, semi-free-ranging and sanctuaries).

174

175

## RESULTS

176 Results are usually presented as proportion or percentages of studies. If studies have  
 177 used multiple species, research environments, methodological approaches, or research foci,  
 178 the sum of these percentages may exceed 100%.

179

180 *Multimodality: Has a larger proportion of primate communication studies been multimodal*  
 181 *in the period 2011-2020 compared to 1960-2008?*

182 Fig. 1 illustrates the proportion of primate communication research published each  
 183 year from 1960-2020. It shows that the number of studies investigating two or more  
 184 modalities in an integrated way in recent years remains low ( $N = 6$ ), with no obvious increase  
 185 in multimodal studies in the decade following the publication of Slocombe et al. (2011).  
 186 Indeed, a Fisher's exact test showed that the proportion of integrated multimodal research  
 187 was significantly lower in the 2011-2020 (6/294) compared to the 1960-2008 period  
 188 (28/553) ( $P = 0.028$ ). Five additional studies used a multimodal approach, but did not  
 189 investigate them in an integrated way, as they either considered two modalities that were  
 190 inherently linked with each other (facial movements produced during vocalizations) or  
 191 because they studied several modalities, but separately from each other.

192

-----

193

Figure 1

194

-----

195

196 *Modalities: Is there an increase in gestural research in monkeys as well as vocal and facial*  
 197 *research in great apes?*

198 The imbalance in the distribution of studies across modalities found in Slocombe et  
 199 al. (2011) was still present in the current dataset including 283 unimodal studies: vocal

200 studies remained the most frequently researched modality (N = 201, 71.0%), and  
201 substantially less research was conducted in the gestural (N = 54, 19.1%) and facial modality  
202 (N = 28, 9.9%).

203       Regarding the consideration of different primate species, a quarter of studies  
204 investigated chimpanzees (N = 72, 24.5%), followed by rhesus macaques (*Macaca mulatta*, N  
205 = 43, 14.6%) and common marmosets (*Callithrix jacchus*, N = 26, 8.8%), with the latter two  
206 most frequently used in neuroscientific studies. When contrasting studies on great apes with  
207 those of other primates (lesser apes, monkeys, prosimians) within each of the three  
208 modalities, in line with Slocombe et al. (2011), we found that great apes were still  
209 differentially represented in research across these three modalities in the latest research  
210 period ( $3 \times 2$  chi<sup>2</sup>-test,  $\chi^2_{(2)} = 48.62$ ,  $P < 0.001$ ).

211       However, Fig. 2 illustrates that – unlike facial studies – vocal and gestural research has  
212 shifted to focus more on their corresponding understudied species in recent years. A  $2 \times 2$   
213  $\chi^2$ -test showed that the proportion of vocal studies with great apes is significantly greater in  
214 the recent period (0.22) compared to the 1960-2008 period (0.09;  $\chi^2_{(1)} = 20.98$ ,  $P < 0.001$ ).  
215 There was a non-significant increase in gestural research with non-great apes (from 0.22 to  
216 0.32;  $\chi^2_{(1)} = 1.32$   $P = 0.251$ ) and a non-significant decrease in the proportion of facial  
217 research with great apes between the original (0.24) and recent period (0.11;  $\chi^2_{(1)} = 2.31$ ,  $P =$   
218 0.128).

219

220

-----

221

Figure 2

222

-----

223

224 *Methodological approaches: Has there been an increase in experimental approaches in facial*  
225 *research and observational methods in vocal studies?*

226 In Slocombe et al. (2011), the proportion of observational and experimental methods  
227 varied significantly across modalities, with vocal studies being the most experimental and  
228 facial expressions the least. For the current dataset, a  $3 \times 2$   $\chi^2$ -test revealed that the  
229 proportion of experimental approaches across modalities varied significantly ( $\chi^2_{(2)} = 13.21$ ,  $P$   
230  $= 0.001$ ). However, the pattern was different to the original period, with the highest  
231 proportion of experimental work found in facial research (0.64) and the lowest in gestural  
232 research (0.24). When each modality was examined individually, the proportion of  
233 experimental approaches to facial expressions increased significantly in the current (0.64)  
234 compared to the original period (0.36;  $\chi^2_{(1)} = 7.48$ ,  $P = 0.006$ ), while for gestures, the  
235 proportion of experimental methods decreased significantly (from 0.49 to 0.24;  $\chi^2_{(1)} = 7.08$ ,  $P$   
236  $= 0.008$ ). In vocal research, the proportion of observational methods increased from 0.47 in  
237 the original period to 0.62 in the recent period ( $\chi^2_{(1)} = 12.48$ ,  $P < 0.001$ ).

238

239 *Research environments: Are there more gestural and facial studies in wild settings?*

240 Slocombe et al. (2011) demonstrated that the majority of research into primate  
241 communication was conducted in captivity, while this pattern was reversed in the most  
242 recent period, with 57% of studies including data from the wild. However, research  
243 environments still differed across modalities in the recent period ( $\chi^2_{(2)} = 21.74$ ,  $P < 0.001$ , Fig.  
244 3). While there was a similar pattern of most studies in wild settings occurring in the vocal  
245 domain, now the least research on wild populations was seen in the facial, not the gestural  
246 domain, as was found in the original period.

247 When each modality was examined individually across the two research periods, there  
 248 was no increase in facial studies in wild settings in the recent period (0.11) compared to the  
 249 original period (0.08;  $\chi^2_{(1)} = 0.18, P = 0.669$ ). In contrast, significantly more research was  
 250 conducted in the wild for both vocal and gestural signals: the proportion of gestural studies  
 251 increased from 0.08 in the original period to 0.46 in the recent period ( $\chi^2_{(1)} = 19.40, P <$   
 252  $0.001$ ), and for vocal research, the proportion of studies in wild settings increased from 0.38  
 253 to 0.57 ( $\chi^2_{(1)} = 18.38, P < 0.001$ ).

254 -----

255 Figure 3

256 -----

257

258 *Research focus: Is there an increased focus on receiver behavior across modalities?*

259 Slocombe et al. (2011) found that the majority of studies investigated signal  
 260 production and producer behavior, although this focus varied across modalities. Examination  
 261 of the recent period indicated that the research focus still varied significantly across  
 262 modalities ( $\chi^2_{(2)} = 8.55, P = 0.014$ ), with most studies examining the receiver found in the  
 263 facial (0.68), then gestural (0.44), and lastly vocal domain (0.39; Fig. 4). When examining  
 264 each modality separately across the research periods, the proportion of vocal studies with  
 265 focus on the receiver remained stable over the two periods (0.38 vs 0.39;  $\chi^2_{(1)} = 0.09, P =$   
 266  $0.761$ ). In contrast, for both facial and gestural research, the proportion of studies  
 267 investigating the perception of these signals and the corresponding receiver behavior  
 268 increased significantly in the recent period compared to the original period (facial: 0.39 to  
 269 0.68;  $\chi^2_{(1)} = 7.49, P = 0.006$ ; gestural: 0.20 to 0.44;  $\chi^2_{(1)} = 7.39, P = 0.007$ ).

270

271

-----

272

Figure 4

273

-----

274

275

276

## DISCUSSION

277

The aim of this paper was to assess the current state of the art in primate

278

communication research and to compare it with the main findings of our Slocombe et al.

279

(2011) paper. In 2011, we found that the different modalities of primate communication

280

were not studied in similar ways, as they each attracted different research questions and

281

methods. We argued that this rendered comparisons across modalities difficult if not

282

inappropriate, with serious implications for theories of language evolution and attempts to

283

identify origins of human language.

284

In the current paper, through systematic review of the primate communication

285

literature from 2011 to 2020, we found that there has been no significant shift in focus or

286

move towards a more multimodal approach. Despite our call for the use of multimodal

287

approaches to study primate communication in more comprehensive and integrated ways,

288

the number of such studies has actually decreased.

289

There are various possible reasons why unimodal primate communication research

290

continues to dominate. First, studying multiple modalities is very challenging and requires

291

training in multiple methods. Researchers and their teams have historically specialized in the

292

theoretical approach and corresponding methods of a single modality and may feel they lack

293

the expertise to incorporate another modality. It takes time to change the historical

294

foundations of a research group (and it is also possible that some scientific funding bodies

295 tend to favor tried and tested approaches, and thus more incremental research). The Open  
296 Science movement might help to push the field towards a more collaborative space, with  
297 shared data and code enabling the adoption of more similar methods across studies. Indeed,  
298 specific projects have been established to promote collaboration and sharing of methods,  
299 which we hope are successful (e.g. Many Primates project: Many Primates et al. 2019,  
300 <https://manyprimates.github.io>, and PhyloPsy: <https://www.phylopsy.org/project>).

301         Second, some methods might simply be more suited to one modality over others and  
302 are difficult to transfer to others. Playback experiments, for example, allow vocal  
303 researchers to explore receiver understanding of signals in captive and wild settings by  
304 simulating group member interactions that are occurring out of sight of the receiver.  
305 However, the corresponding video playbacks necessary to explore understanding of visual  
306 facial and gestural signals would need to be constrained to captive populations. Video  
307 playbacks may also require integration with other measures (e.g., eye tracking) to  
308 extrapolate receiver understanding of third party interactions. Thus, although on a  
309 theoretical level it is a good idea to create consistency between modalities, in practice this  
310 can be very difficult and in some cases impossible (Liebal & Oña 2018).

311         Third, there is still inconsistency in the literature and across disciplines in both how the  
312 term "multimodal" is used, and what makes multimodality interesting. While some scholars,  
313 mostly from the field of comparative psychology, use the term multimodality to refer to  
314 combinations of visual signals (e.g., gesture and facial expression) as well as combinations of  
315 auditory and visual signals (e.g., vocalization and gesture) (Leavens & Hopkins 2005; Pollick  
316 et al. 2007; Liebal et al. 2013; Micheletta et al. 2013; Taglialatela et al. 2015), scholars from  
317 the fields of behavioral ecology and evolutionary biology argue that true multimodality must  
318 combine sensory modalities, not signal types (Partan & Marler 2005; Higham & Hebets

319 2013). There were attempts to integrate these perspectives on multimodality from  
320 comparative psychology and behavioral ecology by using the term “multicomponent” signals  
321 (Micheletta et al. 2013) and to differentiate between bimodal combinations consisting of  
322 two sensory modalities (e.g., visual and acoustic components) as opposed to unimodal  
323 multicomponent signals (consisting of several components of one sensory modality) (Rowe  
324 1999). However, others use the term multicomponent differently and suggest it may be  
325 important to differentiate between the production and perception of multicomponent  
326 communication (Holler & Levinson 2019). With a focus on production, they refer to  
327 multiplex communication if at least two different articulators are involved in producing a  
328 signal, while with regard to perception, they refer to multimodal communication if at least  
329 two different sensory channels are involved (Holler & Levinson 2019). This array of  
330 definitional suggestions demonstrates that the very same term might be defined and  
331 operationalized very differently across disciplines. At the same time, more detailed  
332 terminology has been introduced to try and capture the complexity of multimodal  
333 communication, but we still seem some way off a shared concept. Taken together, the lack  
334 of clear consensus on definitions of multimodal communication, the challenge of developing  
335 expertise and confidence with diverse methodologies required for rigorous research in  
336 multiple modalities, and difficulties in applying some methodologies consistently across  
337 modalities are all likely to have contributed to the low number of multimodal primate  
338 communication studies. More extensive collaboration and open provision of training  
339 sessions targeting methodologies used in vocal, gestural and facial research are likely  
340 needed to assist the field in adopting a more holistic approach to studying communication in  
341 primates.

342           Although there has been no increase in multimodal research over the last 10 years,  
343 some more promising progress has been made in terms of addressing gaps in our knowledge  
344 using unimodal approaches. Considering methodological approaches, there was as a  
345 significant increase in experimental work on facial expressions, as well as a significant  
346 increase in vocal research based on observational approaches. However, while gesture  
347 research had used equal proportions of observational and experimental approaches in the  
348 original period, experimental methods decreased significantly in recent years, which may  
349 need redressing in the coming years.

350           Regarding research environment, there was a shift from an original focus on captive  
351 primates to research on wild populations in the recent period, but this varied across  
352 modalities. Although vocalizations are still the most frequently studied modality in wild  
353 settings, for gesture research, there was a substantial shift from captive to wild settings,  
354 particularly in studies with great apes. However, the gap of knowledge regarding facial  
355 expressions of free-ranging primates still exists, as numbers of such studies remained very  
356 low throughout the 1960-2020 period.

357           Although the research focus still varied across modalities and despite the majority of  
358 research investigating the production side of communication, there was also an increase of  
359 studies considering the perception of signals, especially in the facial and gestural domain.

360           Finally, with regard to the investigated species, the call in Slocombe et al. (2011) for  
361 more vocal research on great ape species seems to have been answered, with a significant  
362 increase in great ape vocal research. However, despite this promising shift, in the 2011-2020  
363 period, there remains a significant difference in the proportion of studies that included great  
364 apes across the three modalities, with the majority of gestural research conducted with  
365 great apes, and the majority of vocal and facial communication focused on non-great ape

366 species. Importantly, almost half of the studies in the current dataset is based on the  
367 investigation of only three primate species (chimpanzees, rhesus macaques, common  
368 marmosets). Thus, although there are often good reasons for studying these species more  
369 than others, it is important to note that our current knowledge about communication across  
370 the primate order is not representative, since the majority of research is based on a very  
371 limited number of primate species.

372         Our recent review of the literature also highlighted two other important issues within  
373 the primate communication field. First, there are still considerable differences in the  
374 research effort dedicated to vocal, gestural and facial research: the majority of research is  
375 still conducted on vocalizations, and despite a recent relative increase in the proportion of  
376 gestural studies, research on modalities other than vocalizations remains scarce. Second,  
377 classifications of signals and the modality to which they belong are not consistent and/or  
378 vary across studies. For example, a signal can be classified as one modality in one study, but  
379 another modality in another study, such as lipsmacks, which are considered facial  
380 expressions, facial gestures, gestures or orofacial movements across studies (Ferrari et al.  
381 2012; Coudé & Ferrari 2018; Clark et al. 2020). The identification of two separate  
382 neuroanatomical routes seems to support the notion of two “types” of facial movements,  
383 which differ in the extent of volitional control (Rinn 1984; Parr et al. 2005). Thus, how these  
384 movements are classified could be important, but nevertheless this differs across studies.  
385 Likewise, primate vocalizations (which use the vocal fold) may or may not be distinguished  
386 from sounds (such as whistles, raspberries), which are made with the mouth but are not  
387 voiced (Leavens et al. 2004; Lameira et al. 2013). Both signal types are auditory, but likely  
388 associated with different physiological and cognitive mechanisms.

389           Taken together, it is promising that some of the gaps in our knowledge highlighted by  
390 Slocombe et al (2011) have started to be addressed: Vocal research in great apes,  
391 experimental approaches in facial research, observational approaches in vocal research, and  
392 studies considering the receiver and signal perception in the gestural and facial domains  
393 have all increased. We hope that these trends continue, but it is important to note that  
394 despite this progress from 2011-2020, we still found significant differences in the  
395 distribution of studies that focus on great-apes, experimental approaches, wild populations  
396 and receiver behavior across the three modalities. In addition, our analyses identified  
397 several outstanding gaps in our knowledge, where no significant progress has been made in  
398 addressing them in recent year. In particular, gestural research on non-great ape species,  
399 facial research on great-ape species, a focus on wild populations in facial research and  
400 greater consideration of the receiver in vocal research, need to be addressed in the years to  
401 come. Considering the current landscape of primate communication findings and  
402 comparative approaches to language evolution, our conclusions are similar to the proposal  
403 made in Slocombe et al. (2011): until we have a more complete picture of primate  
404 communication across modalities and more comparable research results, it is not possible to  
405 reject or support a specific theory of language evolution. Thus, it is important to not  
406 interpret the absence of evidence for a trait in a poorly researched area as an absence of  
407 ability, although many theories of language evolution and many of corresponding studies  
408 present such arguments (Zuberbühler 2005; Tomasello & Call 2019).

409

410

411

## CONCLUSION

412 To conclude, despite increasing theoretical consensus on the importance of a  
413 multimodal approach for studying primate communication in a more comprehensive way,  
414 the vast majority of studies still focus exclusively on either facial expressions, gestures or  
415 vocalizations. Within unimodal approaches to primate communication, many of the  
416 differences in approach and methodology between vocal, gestural and facial research  
417 identified in Slocombe et al (2011) persist: significant differences in the distribution of  
418 studies that include great apes and wild populations and the use of experimental  
419 approaches, as well as the focus on receiver behavior across the three modalities remain.  
420 However, in the last 10 years, significant progress has been made towards addressing some  
421 of the gaps in our knowledge, with more experimental research on facial expressions, more  
422 vocal work with great ape species, and a shift to work on wild primates, particularly in the  
423 gestural domain. Furthermore, human language is increasingly considered multimodal  
424 (Vigliocco et al. 2014; Holler & Levinson 2019). As a consequence, theories are emerging that  
425 propose a multimodal origin of human language (Wacewicz & Zywczyński 2017; Fröhlich et  
426 al. 2019), which provide a new theoretical framework and may further encourage  
427 multimodal approaches in empirical primate communication research.

428

429

#### ACKNOWLEDGEMENTS

430 We thank Federica Amici, Blanca Striegler, and Damilola Olaoba for their support with  
431 coding the data. We are also grateful to Jamie Whitehouse for creating the figures. Many  
432 thanks to the participants of the “Multimodal Communication Workshop” (Erica Cartmill,  
433 Kirsty Graham, Cat Hobaiter, Eithne Kavanagh, Jérôme Micheletta, Linda Oña) for our fruitful  
434 debates on terms and definitions of multimodal research.

435

436

## REFERENCES

- 437 Arbib MA, Liebal K, Pika S. 2008. Primate vocalization, gesture, and the evolution of human  
438 language. *Curr Anthropol.* 49(6):1053-1076. doi:10.1086/593015
- 439 Bergman TJ. 2013. Speech-like vocalized lip-smacking in geladas. *Curr Biol.* 23(7):R268-R269.  
440 doi:10.1016/j-cub.2013.02.038
- 441 Call J, Tomasello M, editors. 2007. The gestural communication of apes and monkeys. New  
442 York (NY): Lawrence Erlbaum Associates.
- 443 Clark PR, Waller BM, Burrows AM, Julle-Danière E, Agil M, Engelhardt A, Micheletta J. 2020.  
444 Morphological variants of silent bared-teeth displays have different social interaction  
445 outcomes in crested macaques (*Macaca nigra*). *Am J Phys Anthropol.* 173(3):411-  
446 422. doi:10.1002/ajpa.24129
- 447 Corballis MC. 2002. From hand to mouth: The origins of language. In: Christiansen MH, Kirby  
448 S, editors. *Language evolution*. Princeton (NJ): Princeton University Press; p. 201-218.
- 449 Coudé G, Ferrari PF. 2018. Reflections on the differential organization of mirror neuron  
450 systems for hand and mouth and their role in the evolution of communication in  
451 primates. *Interact Stud.* 19(1-2): 38-53. doi:10.1075/is.17037.cou
- 452 Fedurek P, Slocombe KE. 2011. Primate vocal communication: a useful tool for  
453 understanding human speech and language evolution? *Hum Biol.* 83(2):153-173.  
454 doi:10.3378/027.083.0202
- 455 Ferrari PF, Vanderwert RE, Paukner A, Bower S, Suomi SJ, Fox NA. 2012. Distinct EEG  
456 amplitude suppression to facial gestures as evidence for a mirror mechanism in  
457 newborn monkeys. *J Cognitive Neurosci.* 24(5): 1165-1172.  
458 doi:10.1162/jocn\_a\_00198

- 459 Fitch WT. 2005. The evolution of language: A comparative review. *Biol Philos.* 20(2):193-203.  
 460 doi:10.1007/s10539-005-5597-1
- 461 Fröhlich M, van Schaik CP. 2018. The function of primate multimodal communication. *Anim*  
 462 *Cogn.* 21(5):619-629. doi:10.1007/s10071-018-1197-8
- 463 Fröhlich M, Sievers C, Townsend SW, Gruber T, van Schaik CP. 2019. Multimodal  
 464 communication and language origins: integrating gestures and vocalizations. *Biol Rev,*  
 465 94(5):1809-1829. doi:10.1111/brv.12535
- 466 Hewes GW. 1992. Primate communication and the gestural origin of language. *Curr*  
 467 *Anthropol.* 33(S1):65-84. doi:10.1086/204019
- 468 Higham JP, Hebets EA. 2013. An introduction to multimodal communication. *Behav Ecol*  
 469 *Sociobiol.* 67(9):1381-1388. doi:10.1007/s00265-013-1590-x
- 470 Holler J, Levinson SC. 2019. Multimodal language processing in human communication.  
 471 *Trends Cogn Sci.* 23(8):639-652. doi:10.1016/j.tics.2019.05.006
- 472 Lameira AR, Hardus ME, Kowalsky B, de Vries H, Spruijt BM, Sterck EH, Shumaker RW, Wich  
 473 SA. 2013. Orangutan (*Pongo spp.*) whistling and implications for the emergence of an  
 474 open-ended call repertoire: a replication and extension. *J Acoust Soc Am.* 134(3):326-  
 475 2335. doi:10.1121/1.4817929
- 476 Leavens DA, Hopkins WD. 2005. Multimodal concomitants of manual gesture by  
 477 chimpanzees (*Pan troglodytes*): influence of food size and distance. *Gesture,* 5(1):75-  
 478 90. doi:10.1075/gest.5.1.07lea
- 479 Leavens DA, Hostetter AB, Wesley MJ, Hopkins WD. 2004. Tactical use of unimodal and  
 480 bimodal communication by chimpanzees, *Pan troglodytes*. *Anim Behav.* 67(3):467-  
 481 476. doi:10.1016/j.anbehav.2003.04.007

- 482 Liebal K, Oña L. 2018. Different approaches to meaning in primate gestural and vocal  
483 communication. *Front Psychol.* 9:478. doi:10.3389/fpsyg.2018.00478
- 484 Liebal K, Waller BM, Burrows AM, Slocombe KE. 2013. *Primate communication: a*  
485 *multimodal approach.* Cambridge (UK): Cambridge University Press.
- 486 Many Primates, Altschul DM, Beran MJ, Bohn M, Call J, DeTroy S, Duguid SJ, Egelkamp CL,  
487 Fichtel C, Fischer J, et al. 2019. Establishing an infrastructure for collaboration in  
488 primate cognition research. *PLoS ONE*, 14(10):e0223675.  
489 doi:10.1371/journal.pone.0223675
- 490 Micheletta J, Engelhardt A, Matthews L, Agil M, Waller BM. 2013. Multicomponent and  
491 multimodal lipsmacking in Crested macaques (*Macaca nigra*). *Am J Primatol.*  
492 5(7):763-773. doi:10.1002/ajp.22105
- 493 Parr LA, Cohen M, de Waal FBM. 2005. Influence of social context on the use of blended and  
494 graded facial displays in chimpanzees. *Int J Primatol.* 26(1):73-103.  
495 doi:10.1007/s10764-005-0724-z
- 496 Partan SR, Marler P. 2005. Issues in the classification of multimodal communication signals.  
497 *Am Nat.* 166(2):231-245. doi:10.1086/431246
- 498 Pereira AS, Kavanagh E, Hobaiter C, Slocombe KE, Lameira AR. 2020. Chimpanzee lip-smacks  
499 confirm primate continuity for speech-rhythm evolution. *Biol Lett.* 16(5):20200232.  
500 doi:10.1098/rsbl.2020.0232
- 501 Pika S, Liebal K, Tomasello M. 2003. Gestural communication in young gorillas (*Gorilla*  
502 *gorilla*): gestural repertoire, learning, and use. *Am J Primatol.* 60(3):95-111.  
503 doi:10.1002/ajp.10097

- 504 Pollick AS, Jeneson A, de Waal FBM. 2007. Gestures and multimodal signaling in bonobos. In:  
505 Furuichi T, Thompson J, editors. The bonobos. Behavior, ecology, and conservation.  
506 New York (NY): Springer.
- 507 Rinn WE. 1984. The neuropsychology of facial expression: a review of the neurological and  
508 psychological mechanisms for producing facial expressions. *PsycholBull.* 95(1):52-77.  
509 doi:10.1037/0033-2909.95.1.52
- 510 Rowe C. 1999. Receiver psychology and the evolution of multicomponent signals. *Anim*  
511 *Behav*, 58(5):921-931. doi:10.1006/anbe.1999.1242
- 512 Seyfarth RM. 2005. Continuities in vocal communication argue against a gestural origin of  
513 language. *Behav Brain Sci.* 28(2):144-145. doi:10.1017/S0140525X05420038
- 514 Seyfarth RM, Cheney DL. 2017. Precursors to language: social cognition and pragmatic  
515 inference in primates. *Psychon B Revi.* 24(1):79-84. doi:10.3758/s13423-016-1059-9
- 516 Singletary B, Tecot S. 2020. Multimodal pair-bond maintenance: a review of signaling across  
517 modalities in pair-bonded nonhuman primates. *Am J Primatol.* 82(3): e23105.  
518 doi:10.1002/ajp.23105
- 519 Slocombe KE, Waller BM, Liebal K. 2011. The language void: The need for multimodality in  
520 primate communication research. *Anim Behav.* 81(5):919-924.  
521 doi:10.1016/j.anbehav.2011.02.002
- 522 Taglialatela JP, Russell JL, Pope SM, Morton T, Bogart S, Reamer, LA, Shapiro SJ, Hopkins WD.  
523 2015. Multimodal communication in chimpanzees. *Am J Primatol.* 77(11):1143-1148.  
524 doi:10.1002/ajp.22449
- 525 Tomasello M. 2008. *Origins of human communication.* Cambridge (MA): The MIT Press.
- 526 Tomasello M, Call J. 2019. Thirty years of great ape gestures. *Anim Cogn.* 22(4):461-469.  
527 doi:10.1007/s10071-018-1167-1

- 528 Vigliocco G, Perniss P, Vinson D. 2014. Language as a multimodal phenomenon: implications  
529 for language learning, processing and evolution. *Philos T Roy Soc B*. 369: 20130292.  
530 doi:10.1098/rstb.2013.0292
- 531 Wacewicz S, Zywczyński P. 2017. The multimodal origins of linguistic communication, *Lang*  
532 *Comm*. 54:1-8. doi:10.1016/j.langcom.2016.10.001
- 533 Zuberbühler K. 2005. The phylogenetic roots of language: evidence from primate  
534 communication and cognition. *Curr Dir Psychol Sci*. 14(3):126-130.  
535 doi:10.1111/j.0963-7214.2005.00357.x
- 536 Zuberbühler K. 2015. Linguistic capacity of non-human animals. *Wires Cogn Sci*. (3):313-321.  
537 doi:10.1002/wcs.1338
- 538
- 539

540 FIGURE CAPTIONS

541

542

543 Fig. 1. — Number of unimodal (vocal, facial and gestural) and multimodal studies published

544 between 1960 and 2020.

545

546

547 Fig. 2. — Illustration of the percentage of studies in each time period within vocal, gestural  
548 and facial research that included great ape species (light grey), or focused exclusively on  
549 non-great ape species (dark grey: monkeys, hylobatids, prosimians). In the 1960-2008  
550 period, the number of studies reported were N = 122 for facial, N = 51 for gestural, N = 352  
551 for vocal (Slocombe et al., 2011). In the 2011-2020 period, the number of studies reported  
552 were N = 28 for facial, N = 54 for gestural, N = 201 for vocal.

553 Fig. 3. — Percentages of studies conducted in wild and captive settings, shown for each  
554 modality for the previous (1960-2008) and the current period (2011-2020).  
555

556 Fig. 4. — Percentages of studies investigating the production (signaler behavior) and  
557 perception (receiver behavior) in the periods of 1960-2008 and 2011-2020.