

1 **Language as a tool for social bonding: evidence from wild chimpanzee gestural, vocal**
2 **and bimodal signals**

3 Ljubica Damjanovic¹, Sam G. B. Roberts¹, Anna I. Roberts²

4
5 ¹ School of Psychology, Liverpool John Moores University, Liverpool, L3 3AF,
6 UK

7 ² Institute of Human Biology and Evolution, Adam Mickiewicz University,
8 Poznan, Poland

9 Correspondence: anna.roberts@amu.edu.pl

10 Key words: chimpanzees, gestural communication, visual attention, social
11 network, intentionality

12
13
14 Final publication

15 Damjanovic, L., Roberts, S. G., & Roberts, A. I. (2022). Language as a tool for
16 social bonding: Evidence from wild chimpanzee gestural, vocal and bimodal
17 signals. *Philosophical Transactions of the Royal Society B*, 377, 20210311.

18 <http://doi.org/10.1098/rstb.2021.0311>

19

20

21

22

23 **Language as a tool for social bonding: evidence from wild chimpanzee gestural, vocal and** 24 **bimodal signals**

25 **Abstract**

26 The evolution of language has fascinated anthropologists, psychologists and biologists for
27 centuries, seeking to infer language origins from the communication of primates, our closest
28 living relatives. Capacity for intentional signalling is a key feature of transition to language in
29 our hominin ancestors, facilitating complex social dynamics in complex social groups.
30 However whether vocal, gestural and bimodal signals are differentiated according to
31 intentional use and hence complex sociality has not been studied, making unclear the modality
32 of language evolution. We addressed this question in wild chimpanzees. We found that larger
33 social network size was associated with a larger network of gestural but not vocal or bimodal
34 signals. Response waiting was more common in association with gestures than vocalisations,
35 but elaborations were more common in vocal than gestural or bimodal signals. Overall,
36 chimpanzees were more likely to manage weak social bonds through vocalisations, whereas
37 strong social bonds were managed through gestures and bimodal signals. However, when social
38 bonds were weak, gestures accompanied by response waiting were more likely to elicit
39 approaches than vocalisations accompanied by elaboration, which elicited avoidance. This
40 suggests that gestures were the primary modality of language evolution and that the use of more
41 sophisticated gestural signalling led to evolution of complex social groups of hominin
42 ancestors.

43 **1. Introduction**

44 Understanding whether language evolved to increase group cohesion and facilitate life
45 in complex social groups is one of the most important questions in establishing the evolutionary
46 pressures driving evolution of human social cognition [1, 2]. In order to determine origins of

47 the language evolution, studies have primarily focused on examining communication of
48 primates. Producing sounds by means of a vocal tract, is a primary characteristic of language,
49 and therefore the bulk of studies of primate communication, have focused on vocal signals [3].
50 Vocalizations are functionally referential, and therefore like language can convey semantic
51 meanings to the recipients, suggesting that language evolved from vocal modality of
52 communication [4]. Further, it was argued that the combined vocalization-gesture system
53 (conventionally referred to as bimodal communication) may have driven the evolution of
54 language, whereby gesture and vocalization had different but complementary functions [5].
55 According to this hypothesis, language evolved from bimodal gesture-vocal signals because it
56 can identify greater number of distinct messages than either vocal or gestural signals alone, and
57 therefore like language can increase the complexity of messages that can be conveyed by
58 communication [6].

59 However, recent theoretical approaches to language evolution posit that language
60 evolved not from vocalizations or bimodal signals, but initially from gestures. Gestural
61 communication, defined as socially directed movements of the head, body, hands or the body
62 postures is an important ancestral feature of language evolution that humans share with their
63 primate relatives [7]. Human language is underpinned by neurological structures that are
64 homologous to the ones underpinning gestural communication [8]. Only humans and apes, use
65 gestural communication as a primary modality of signaling, which was shown to be more
66 flexibly deployed in social interactions than either facial expressions or vocalizations [9]. Here
67 we examine whether the predecessors to language evolution can be found in wild chimpanzee
68 gestural, vocal or bimodal signals.

69 In order to understand the homology in complexity of communication between humans'
70 and apes' it is important to examine understanding of intentionality in signaling, defined as
71 ability to appreciate that others have different thoughts from us, and that these thoughts affect

72 their behavior [10]. In intentional signalling, the signaller has a goal and flexibly modifies
73 communication to influence the comprehension state of the recipient [10]. This ability is one
74 of the most important features of human communication and makes social relationships of
75 humans so complex. The most important question is whether primates have understanding of
76 intentionality in gestures, or if vocal and bimodal signals are also characterized by
77 understanding of intentionality.

78 Studies of gesturing in captive chimpanzees have shown that they have some
79 understanding of intentionality as evidenced by a signaler's audience checking (directing visual
80 attention at recipient prior to signal), response waiting (directing visual attention at recipient
81 after signaling) and elaboration of signals (using a new signal after the first signal in sequence)
82 until their goal is obtained, or failure is indicated [11]. For example, chimpanzees can figure
83 out if an experimenter knows, or does not know about the location of a hidden food and persist
84 in gesturing until the experimenter finds the food [12]. Some studies have proposed that human
85 contact is necessary for chimpanzees to have flexible, intentional communication [13]. This is
86 because humans interact with chimpanzees in different ways than chimpanzees do amongst
87 themselves, for instance, by attempting to direct their attention towards self or other objects or
88 events. When humans display these behaviors toward captive chimpanzees, the apes may
89 acquire different abilities to those of wild chimpanzees, that is, the enculturated apes may have
90 specifically adapted to contact with humans [14]. However, research into intentionality
91 underlying gestures in wild chimpanzees shows that intentional gestures are more ubiquitous
92 than previously thought. Chimpanzees in the wild show complex use of gestures demonstrated
93 by how chimpanzees make *sequences* of gestures, i.e. gestures produced consecutively, in
94 response to another's behavior [15]. Chimpanzees show an awareness of whether or not the
95 recipient of the gesture really understands the message the sender is trying to get across. If the
96 recipient only partly understands the message, then the sender repeats the same gesture, while

97 if the recipient does not at all understand the message, then the sender uses new gesture. These
98 findings are significant because they show that gestures are not simply a result of emotional
99 states felt by a chimpanzee but are choices designed to influence others, in order to achieve
100 desired goal. Intentional gesture is not acquired in captivity, but is a trait shared with humans
101 and other apes [13].

102 It is not only the gestural behavior that can shed light on our evolutionary origins; the
103 vocal and bimodal signals are also important in 'decoding' communicative and cognitive
104 abilities [16]. To address understanding of intentionality in bimodal signals, Leavens & authors
105 [17] showed that captive chimpanzees tailor use of bimodal and unimodal signals to success of
106 communication acts with a human observer. Specifically designing playback experiments to
107 capture the nature of intentional communication in alarm calls of wild chimpanzees, Schel &
108 colleagues [18] showed that alarm calls experimentally elicited by a presence of a snake were
109 socially directed, and used in tandem with audience checking and gaze alternations until the
110 recipients were safe from the snake. These findings thus suggest, that chimpanzees are unique
111 in having in common with us the propensity to use vocal and bimodal signals to alter the
112 knowledge states of the recipient, which is a key feature important in the evolution of human
113 communication [8, 18].

114 However, understanding intentionality in primate communication has been complicated
115 by the methodological constraints of focusing on the signaler's perspective, whereas recent
116 studies point out the important role of intentionality from the point of view of the recipient.
117 According to this perspective intentional signals influence knowledge states of the recipient
118 because they augment the recipient's capacity to process information about social relationships,
119 whereby the recipient focuses on relevant information about ongoing and past social
120 interactions to figure out the signaler's goal. Intentionality in communication is required to
121 form social bonds in socially complex species. However, absence of previous experience with

122 the social partner may limit the ability of the recipient to process information about social
123 relationships. Weak social bonds are ambiguous, and the high unpredictability of social
124 interactions causes the recipient to perceive them as threat resulting in high likelihood of
125 avoidance [19-21]. In this context, higher intensity signals attract attention because they have
126 a high predictive value for the presence of threat, causing reflexive avoidance in response to
127 physiological stimulus of physical characteristics [22]. In contrast, intentionality in
128 communication can increase the positive evaluation of the social interaction, facilitating
129 approaches and enabling two animals to engage in social interaction. This process gives rise to
130 efficient social interactions between strongly bonded dyad partners, based on simple readouts
131 of behavioral states, whereby recipients learn to respond to emotional state generated as a
132 consequence of direct, repeated experience of the social interaction with the signaller.

133 Chimpanzees live in complex fission-fusion social system, whereby they split into
134 subgroups of different composition and duration on a daily basis, leading to a presence of weak
135 social bonds. Although the size of the chimpanzee communities on average is very large,
136 ranging from 20-150, individual chimpanzees differ in the size of their social networks, with
137 central chimpanzees in the communities having a large number of strong social connections,
138 whereas peripheral chimpanzees have smaller numbers of strong social connections. An
139 important avenue of research is to use social network analyses to examine how the social
140 relationships of primates and intentionality in communication are inter-connected. Such
141 analyses will reveal how primates use communication and other behaviours to facilitate life in
142 complex social groups and whether primate communication was a predecessor to language
143 evolution.

144 Here we examine this topic in wild chimpanzees and hypothesize that gestures are more
145 effective in social bonding than either vocal or bimodal signals because they are more
146 intentional. Chimpanzees (*Pan troglodytes schweinfurthii*) are among our closest living

147 relatives and examining the association between communicative and social complexity in this
148 species is particularly interesting because of the complexity of their fluid fission-fusion social
149 system [23]. They are thus excellent model species to investigate patterns of sociality and
150 communicative complexity so as to inform models of human social evolution. To this end we
151 examined whether: 1) the complexity of the gestural, vocal and bimodal communication
152 network is positively associated with the complexity of the social network; 2) gestural, vocal
153 and bimodal communication differ in respect of 3) intentionality, 4) strength of social bond and
154 5) how weak social bonds are managed, and 6) social interactions are differentiated according
155 to the rate of gestural, vocal and bimodal communication.

156 **2. Methods**

157 (a) Study site, data collection and coding

158 Habituated adult chimpanzees (6 males, 6 females) were followed over 9 month period at the
159 Budongo Conservation Field Station, Budongo Forest Reserve in Uganda (2006 – 2008). The
160 observation duration for each subject is shown in Supplementary Information 1 [24]. Focal
161 follows of 18-minute duration (9 scans at 2-minute intervals) recorded the activity of the focal
162 individual; the identity, activity, and distance of the nearest adult neighbor relative to the focal
163 subject as well as identity of all chimpanzees present within 10 m of the focal subject
164 continuously recording communication using a digital video camera. Supplementary
165 Information 1 gives description of the categories of behaviour coded [24]. For each
166 communication event, we determined the social bond of the recipient towards the signaller
167 from 8 social behaviours using the Composite Sociality Index [12] (Supplementary
168 Information 1) [24].

169 (b) Social network analysis

170 We deployed Double Dekker Semi-Partialling Multiple Regression Quadratic Assignment
171 Procedure (MRQAP) to examine the associations between communication networks,
172 calculated as the frequency of behaviour per hour dyad partners spent within 10 meters [13]
173 and the social networks calculated as the frequency of behaviour, when nearest neighbor,
174 within 2 m per hour spent within 10 m. We considered six behaviors were both focal and non-
175 focal subject were engaged in the same social behavior with each other (visual attention,
176 proximity, feeding, resting travel, mutual grooming) and two behaviours which were
177 unidirectional (giving and receiving grooming). The data entry and the results of these analyses
178 are reported in Supplementary Information 1 [24]. Using matrices of social interactions, we
179 calculated centrality measures using normalized degree centrality. This measure represents the
180 average value of each row or column of the network matrix (i.e., the average value of that
181 behavior for each focal chimpanzee). Since the network of social behaviors differed between
182 dyads A to B and B to A, we calculated indegree (behaviors directed by conspecifics toward
183 the focal chimpanzee) and outdegree (behaviors directed by the focal chimpanzee to
184 conspecifics) separately. Further, we dichotomized and symmetrized social and
185 communication networks, to obtain the measure of overall network size (the total number of
186 edges connected to a particular node: n degree), we calculated the normalized degree (n degree)
187 of social and communication networks, dichotomizing and symmetrizing social networks. For
188 analyses of social interactions we first obtained composite sociality index using formula
189 presented in Supplementary Information. Both A to B and B to A dyads were considered.
190 Further, we computed the network of strong and weak bonds separately, using composite
191 sociality index, whereby values above 1 represent dyads that have higher rate of social
192 interactions per hour spent within 10 meters than average dyad (strong bond). Values below
193 mean of 1 denoted dyads that have a lower rate of social bonding behavior per hour spent within
194 10 meters than average dyad (weak bond). To construct communication networks, we used

195 rates of communication produced by the focal towards non-focal chimpanzee. In all analysis
196 of centrality in social network, we included four control variables: proximity to oestrous female
197 outdegree (duration of time focal subject spent in proximity to oestrous female per hour spent
198 in the same party outdegree), proximity to kin outdegree (duration of time focal subject spent
199 in proximity to kin per hour spent in the same party outdegree), sex (two levels: male, female),
200 age (age of focal subject in years) (see Supplementary Information 2 for details of all
201 models)[24]. UCINET 6 for Windows was used to carry out all data transformations and social
202 network analyses.

203 (c) Generalized linear mixed models (GLMM)

204 In the analysis, only those events were included which occurred as first signal in the sequence.
205 We included all behaviors potentially communicative (e.g. quadrupedal stance), unlike in our
206 previous studies [25], excluding only the scratch behavior. This was done so as to take into
207 account the influence of behaviors on social bonding that appeared to be communicative but
208 were potentially non-intentional. To examine differences between gestures, vocalisations and
209 bimodal signals, in the analysis we included five control predictors: signaller and recipient age
210 (two levels: subadult, who was individual less than 15 years old, adult, all others) signaller and
211 recipient sex (two levels: female, male), relatedness (two levels: non-kin, such as mother
212 offspring dyads or siblings, father offspring dyads were largely unknown; kin). In those
213 analyses where social bond of the recipient towards signaler was included, only adult recipients
214 were considered. In the GLMM, three key variables of interest were: 1) communication type:
215 vocalization versus gesture (two levels: vocalization, gesture), bimodal signal versus gesture
216 (two levels: bimodal signal, gesture), bimodal signal versus vocalization (two levels: bimodal
217 signal, vocalization); 2) intentionality of signaling: audience checking (two levels: absent,
218 present), response waiting (two levels: absent, present), elaboration (two levels: absent,
219 present); 3) recipient/ signaller bond (two levels: weak, strong); 4) recipient's response (two

220 levels: avoidance, approach); activity change (two levels: maintain ongoing activity, initiate
221 activity change by directing movement and attention of the recipient). The data in GLMM, had
222 a hierarchical structure: Level 1 (identity of signaller), Level 2 (identity of recipient of the
223 gesture).The models were fitted using a binomial error structure with logit link. Signaller
224 identity and signaller identity by recipient identity, were random effects for which random
225 intercepts were used (see Supplementary Information 2 for details of all models not given in
226 text of the manuscript) [24]. The results of predictors of intentionality by communication type
227 are reported in Supplementary Information 1 [24]. We analyzed all data using SPSS 25.0 (SPSS
228 Inc., Chicago, IL, USA). The datasets are available on Figshare [24].

229 **3. Results**

230 **Overview of the communication events**

231 Overall, we extracted from the footage 5328 potentially communicative behaviours. Out of
232 these cases we excluded 1120 events which contained scratches or combinations of scratches
233 with communication. The remaining dataset contained 3275 events of gestures, 509 bimodal
234 combinations of gestures with vocalisations and 424 events of vocalisations. We coded 3046
235 of communication events for audience checking, 2979 events for response waiting, 3402 events
236 for elaboration, 3206 events for recipient's response and 3401 events for activity change. See
237 Table 1 for presence and absence of these markers according to communication type.

238 **Does complexity of communication network increase with social complexity?**

239 Across all adult dyads in the community where there was some level of social interaction
240 (values of CSI were above 0), the chimpanzees had a differentiated number of social bonds and
241 this varied across two consecutive years. Of the total community size of 36 adult members in
242 2007, and 30 adult members in 2008, the mean \pm sd number of all connections was 4.66 ± 3.67
243 and 9.08 ± 2.42 connections in 2007 and 2008 respectively, and this difference was statistically

244 significant (Wilcoxon signed-rank test, $T = 28$, $z = -2.375$, $p = 0.016$, median 2007: 4, median
245 2008: 8.5). The mean \pm sd number of strong connections in 2008 (6.75 ± 2.83 connections)
246 was significantly higher than in 2007 (4 ± 3.31 connections): (Wilcoxon signed-rank test, $T =$
247 21 , $z = -2.214$, $p = 0.031$; median 2008: 7.5, median 2007: 3). There was no difference in the
248 number of weak social bonds in 2007 (the mean \pm sd: 0.66 ± 1 connections) and 2008 (the
249 mean \pm sd; 2.33 ± 1.49 connections): (Wilcoxon signed-rank test, $T = 19.50$, $z = -1.897$, $p =$
250 0.094 ; median 2007: 0, 2008: 2.5) (see Supplementary Information for number of strong and
251 weak social bonds per each focal chimpanzee across the consecutive years). Across 132 dyads
252 of focal chimpanzees, animals directed communication at a dyad partner at a mean rate \pm sd of
253 3.96 ± 11.06 gestures, 0.47 ± 4.79 vocalisations and 1.08 ± 2.87 bimodal signals. Chimpanzees
254 used vocal, gestural and bimodal signals with the majority of their network members. The mean
255 degree of gestures was 53% of connections to all network members, bimodal signals was 37.9%
256 of connections and vocalisations was 12.1% of connections. We used node level regressions to
257 examine whether centrality in the social network predicted centrality in the gesture, bimodal
258 and vocalisation networks. We found that there was a significant positive association between
259 social network size (composite sociality index n degree) and the size of the network of gestural
260 communication ($r^2=0.861$, $\beta= 0.680$, $p = 0.049$), but not bimodal signals ($r^2=0.669$, $\beta= 0.576$,
261 $p = 0.099$) or vocalisations ($r^2=0.706$, $\beta= 0.090$, $p = 0.407$) (Fig. 1).

262 Considering size of strong and weak social bonds networks separately, we found that size of
263 strong social bond network positively predicted size of gesture ($r^2=0.873$, $\beta= 0.803$, $p = 0.040$)
264 and bimodal signal ($r^2=0.808$, $\beta= 0.896$, $p = 0.028$) networks but not size of the vocal network
265 ($r^2=0.803$, $\beta= 0.514$, $p = 0.149$). In contrast, size of weak social bond network negatively
266 predicted size of vocal gesture network ($r^2=0.903$, $\beta= -0.725$, $p = 0.049$) but not the size of
267 gesture ($r^2=0.667$, $\beta= -0.325$, $p = 0.246$) or bimodal signal ($r^2=0.595$, $\beta= -0.498$, $p = 0.145$)
268 networks. Further, we used node level regressions to examine whether centrality in the social

269 network based on social behaviour directed by non-focal subjects at the focal subject
270 (composite sociality index indegree) was predicted by vocal, gestural or bimodal
271 communication produced (communication outdegree) and received (communication indegree).
272 We found that chimpanzees who received a higher rate of social behaviour also received
273 gestural communication (gesture indegree: $r^2=0.753$, $\beta= 0.661$, $p = 0.038$) and bimodal
274 communication (bimodal signal indegree: $r^2=0.745$, $\beta= 0.804$, $p = 0.027$) at higher rates than
275 the chimpanzees who received a lower rate of social behaviour (composite sociality index
276 indegree). However, chimpanzees who received social interactions at a higher rate did not
277 receive vocalizations (vocalisation indegree: $r^2=0.837$, $\beta= 0.713$, $p = 0.064$) at a higher rate
278 than the peripheral chimpanzees in the social network.

279 **Do signallers differentiate use of gestures, bimodal signals and vocalisations according to**
280 **strength of social bond with the partner and use of intentionality markers?**

281 We used GLMM to examine the association between type of communication (vocal versus
282 gestural, bimodal versus gestural, bimodal versus vocal) and the following predictors: strength
283 of social bond of recipient towards signaller, absence and presence of intentionality marker
284 considered separately (audience checking, response waiting, elaboration), controlling for
285 signaller age, relatedness, signaller and recipient sex (Table 2).

286 Comparing vocal and gestural signals, we found that subadult chimpanzees were more likely
287 to direct gestures at the partner than the adults, gestures were more likely used to communicate
288 with kin than non-kin, and female than male recipient. Further, gestures were more likely
289 directed at strongly bonded recipient to the signaller when compared with the weakly bonded
290 recipient. Further, gestures were less likely to be accompanied by elaboration than
291 vocalisations.

292 When comparing bimodal signals to gestures, the analysis showed that gestures were more
293 likely produced by subadults when compared with adults, by males when compared to females,
294 and towards female recipients than to male recipients.

295 Finally, comparing bimodal signals to vocalisations, we found that bimodal signals were more
296 likely when the social bond of the recipient to the signaller was strong than when weak. We
297 also found that bimodal signals were less likely to be accompanied by elaboration than were
298 vocalisations.

299 **Does vocal, gestural and bimodal communication differ according to how it manages**
300 **weak social bonds?**

301 We used GLMM to examine whether approach or avoidance was a more likely response to
302 gestures between weakly bonded dyad partners in response to presence and absence of an
303 intentionality marker (audience checking, response waiting, elaboration), while controlling for
304 signaller age, relatedness, and the sex of signaller and recipient (Table 3). Comparing gestures
305 to vocalisations, we found that when social bonds were weak, communication was more likely
306 to elicit approach to than avoidance from kin compared to non-kin, when communication was
307 accompanied by response waiting, when there was absence of elaboration, and when
308 communication was gestural rather than vocal (Fig. 2). Comparing gestures to bimodal signals
309 we found that approach was more likely than avoidance when communication was
310 accompanied by presence of response waiting and absence of elaboration and when
311 communication was gestural than bimodal.

312 **Do audience checking, response waiting and elaboration between weakly bonded**
313 **partners have a different effect on the recipient according to modality of communication**
314 **used?**

315 We used GLMM to determine whether approach or avoidance to communication accompanied
316 by either audience checking or response waiting in social interactions between weakly bonded
317 dyads occurred in response to vocal versus gestural, and in response to bimodal versus gestural
318 signals, controlling for signaller age, relatedness, signaller and recipient sex. Comparing vocal
319 and gestural signals, we found that chimpanzees were more likely to approach in response to
320 communication made by kin than non-kin ($\beta = -14.222 \pm 0.858$, $p < 0.001$), when the recipient
321 of communication was a male than a female ($\beta = -1.540 \pm 0.694$, $p = 0.028$), and when
322 communication was gestural than vocal ($\beta = -3.690 \pm 0.832$, $p < 0.001$). Further, comparing
323 bimodal and gestural communication, we found that approaches were more likely by male than
324 female recipients ($\beta = -1.501 \pm 0.704$, $p = 0.034$) and when communication was gestural than
325 bimodal ($\beta = -0.974 \pm 0.403$, $p = 0.017$).

326 Next, we used GLMM to examine whether approach or avoidance to the last gesture in an
327 elaboration sequences between weakly bonded dyad partners occurred in response to gestural
328 compared to vocal signals, controlling for signaller age, relatedness, signaller and recipient sex.
329 We found that approach was more likely than avoidance when elaboration was produced by
330 female than a male ($\beta = 14.859 \pm 1.191$, $p < 0.001$) and when elaboration was gestural rather
331 than vocal ($\beta = -0.666 \pm 0.050$, $p < 0.001$).

332 **4. Discussion**

333 In this study, we hypothesized that gestures are more intentional than either vocal or
334 bimodal signals and hence facilitate more complex social dynamics acting as a primary
335 modality of language evolution. We deployed social network analysis to examine the
336 associations between size of the social network and the size of the network of vocal, gestural
337 and bimodal signals. We found that chimpanzees who had a larger number of social
338 connections also had a larger number of connections maintained through gestures but not

339 through vocal or bimodal signals. Using GLMM to examine this association, we found that
340 chimpanzees directed elaborations and vocal signals at weak bonds, whereas they directed
341 gestural and bimodal signals at strong bonds in absence of elaboration. Finally, to examine the
342 effect of signals on the recipient, we found that when social bonds were weak intentional
343 gestures elicited approaches, whereas intentional vocal and bimodal signals elicited avoidance.
344 These data suggest that gestural communication is underpinned by understanding of intentions,
345 whereas vocalisations and bimodal signals are underpinned simply by understanding of
346 behaviour, suggesting that the capacity for intentional signalling in gestures facilitates more
347 complex social dynamics of wild chimpanzees. On this basis we conclude that language as a
348 tool for social bonding has primarily evolved from gestures to facilitate group cohesion in large
349 and complex social groups of hominins.

350 If intentionality is important for managing weak, rarely reinforced relationships, then
351 chimpanzees should direct communication that requires only behavioural understanding at
352 individuals with whom they have strong social bonds and communication underpinned by
353 intentionality at those with whom they have weak bonds. We found that gestural and bimodal
354 signals were more likely in association with strong social bonds, whereas vocalisations
355 primarily co-occurred with weak social bonds. The observed results are likely because gestural
356 signals differ from vocal signals in the cognitive skills that need to be employed in processing
357 of information due to the differences in intentionality. Forming social relationships in novel
358 social conditions, such as interactions with weakly bonded dyad partners, requires the ability
359 of the recipient to attribute goals to social interactions, by allocating memory to selectively
360 focus on relevant information [26]. In contrast, direct social experience reinforces strong social
361 bonds, facilitating understanding of behaviour. If vocalisations are primarily associated with
362 weak social bonds, this would suggest that they are key for forming social bonds with group
363 members. However, attention allocation involves frequent errors and this implies that it is less

364 likely to be successful, biasing strategy towards less cognitively complex tasks that produce a
365 higher chance of a successful outcome [27]. In this context, vocalisations are less cognitively
366 demanding because they influence understanding of behaviour between weak social bonds,
367 whereas gestures are more cognitively complex because they influence understanding of
368 intentions between strong social bonds [28]. Although vocalizations appear to be more
369 intentional than gestures because chimpanzees elaborated with vocalisations more often
370 towards weak social bonds, these elaborations are likely expressions of emotional state of the
371 signaller, rather than attempts to inform the unaware recipient of the goal of the interaction.
372 For instance, when social bonds were weak only gestures accompanied by response waiting
373 elicited approaches, whereas elaborations by vocalisations elicited avoidance. This could
374 potentially indicate that elaborations of vocal signals were threats that were used to manage
375 spatial positions in competitive contexts.

376 Although we acknowledge that not all vocalisations may function in this way (e.g.
377 contact grunts of baboons may not [29]), in the context of our data, vocalisations appear to
378 express signaller's own emotional arousal, and cause the recipient to experience it through the
379 vocal communication of a signaller, as a variation in the state of physiological activation,
380 indicated by heart rate [30], cortisol secretion [31] or nasal temperature [32]. High intensity
381 vocalisations (e.g. alarm call) influence recipient's behaviour by increasing their heart rate and
382 cortisol. The role of vocalisations in maintaining weak social bonds rather than forming new
383 social connections is supported by the data showing that when social bonds were weak
384 chimpanzees used gestures to initiate change of activity by the recipient, whereas vocalisations
385 maintained ongoing activities. The fact that only vocalisations and not gestures or bimodal
386 signals co-occurred with a longer duration of time spent feeding in close proximity, supports
387 this suggestion. This is further supported by the results showing that the size of the social
388 network increased in association with the size of the network of gestural but not vocal signals.

389 One interesting finding was that when weakly bonded dyad partners interacted, the
390 intentionality markers had a different effect on the recipient, depending on the type of
391 communication. Gestures accompanied by intentionality markers elicited approach, whereas
392 with vocalisations accompanied by intentionality markers elicited avoidance. Social
393 environments impose chronic and acute stress on processing of information. It is expected that
394 when these challenges reach a certain threshold, the recipients will downgrade understanding
395 of intention to understanding of behaviour. Acute stress has a larger detrimental effect on
396 understanding of intention than chronic stress, in that in conditions of acute stress,
397 understanding of behaviour is favoured over understanding of intention [33]. Visual attention,
398 such as visually monitoring the dyad partners prior to communication (i.e. audience checking)
399 or after communication (i.e. response waiting) is processed in the areas of the brain involved
400 in both understanding of intention as well as understanding of behaviour [33]. Visual attention
401 during chronic stress disinhibits understanding of intentions, leading to a greater capacity of
402 the recipient to process information. In contrast, visual attention heightens acute stress, leading
403 to a greater inhibition of intentional state of the recipient, and increased behavioural processing.
404 In the context of managing weak social bonds, animals may exploit their repertoire of
405 vocalizations to match that of an interaction partner to whom they actively pay attention to in
406 a stimulus driven way. Use of intentionality markers in this context leads to a greater degree of
407 overlap in the repertoire of vocalisations through enhanced ability of the dyad partners to
408 actively match and synchronise their communication in condition of acute stress. If the
409 motivational state of fear is reduced by a joint display of affect in response to behavioural
410 convergence in vocalisations, the use of intentionality markers facilitates formation of social
411 bonds by automatic attribution of value of 'liking' and 'wanting' [34, 35].

412 It could be argued that language evolution occurred in vocal communication because
413 our study shows that chimpanzees used vocal signals to manage social relationships with

414 weaker links, whereas gestures were used to manage strong social bonds. However, the value
415 attribution in language occurs through cognitively complex processing. For instance, activation
416 of brain areas such as the superior and middle temporal gyri in the left hemisphere which
417 indicate abstract higher-level processes are present during normal language processing, but are
418 reduced when speech sounds are artificially distorted to disrupt the connection between sound
419 and meaning [36]. Gestural communication facilitates formation of the representations about
420 goal of the social interaction with weak social bonds, suggesting a potential important role of
421 these behaviors in language evolution. Psychopharmacological studies examining the influence
422 of dopamine on language processes showed that dopamine enhanced value encoding in speech
423 [37]. This suggests that cognitive processing of primate gestures might be affected in a similar
424 manner to language processing. Cognitive skills underpinning gestural communication
425 therefore probably preceded language evolution.

426 However, the claim for gestural origins for language confounds language capacity
427 with language expression (and has no explanation for why it should have switched from
428 gestures to voice in humans). Humans use speech/language not to manage close relationships
429 (strong ties, which we deal with mainly by gestures) but to manage distant weak ones, just as
430 our data suggest is the case of chimpanzees. Language as an adjunct to grooming to allow
431 group size to increase beyond the limit at ~50 that we can manage via grooming. Although
432 cognitive skills underpinning language, may have first evolved in gestural modality of
433 communication, the language evolution may have occurred in the modality of the low
434 intensity vocal signals once our human ancestors gained greater voluntary control over their
435 vocal output to face challenges of acute stress. As familiar dyad partners repeatedly interacted
436 in a goal directed way through novel, low intensity vocalisations, this innovation could
437 produce group level differences in communication where small differences in morphology of
438 the same type of low intensity vocalisation would produce ‘accents’. Such vocalisations

439 acquired with familiar conspecifics may have been used with unfamiliar dyad partners to
440 establish social bonds when exposed to acute stress as the dopamine is not released in
441 response to habitual signals and the low intensity of signal would prevent overstimulation of
442 the recipient and allow recognition of social distance between the signaller and the recipient
443 when social bonds were weaker. This would allow transfer of value in non-aversive way
444 during acute stress, for instance, signalling value information to the recipient about signaller's
445 potential as a grooming partner giving rise to language evolution.

446 Through hominin evolution there has been an increase in both brain size and this is
447 likely to have been accompanied by an increase in group size [1]. Dunbar [35] has argued that
448 the pressure to maintain larger social groups through hominin evolution may have driven the
449 evolution of language as a novel social bonding mechanism that is more time efficient than
450 grooming. Between primate species, it has been shown that evolutionary increases in the size
451 of the vocal repertoire in non-human primates were associated with increases in both group
452 size and also time spent grooming [38]. This suggests that vocal communication may indeed
453 play a key role in the evolution of social behaviour - larger groups are more complex to manage,
454 and thus require a larger repertoire to maintain an increasing number of differentiated
455 relationships. However, it is increasingly being recognised that gestural communication also
456 plays a key role in regulating social behaviour, and systematic studies of the role of gestural
457 communication in wild primates are in their infancy. This study examined how the gestures
458 and vocalisations and bimodal signals in wild chimpanzees are related to the complexity of
459 social network. There is currently an active debate as to whether human language evolved from
460 vocal or gestural communication, and the finding that the use of gestural but not bimodal or
461 vocal communication is related to a greater complexity of social network provides important
462 insights into this debate. A key challenge of the future study of primate sociality is evaluating

463 the relative importance of vocalisations, bimodal signals and gestures in the maintenance of
464 primate social networks [39], and exploring how primates in groups of increasing size use these
465 behaviours differentially to maintain their social relationships.

466 **Acknowledgments**

467 National Science Centre, Poland: UMO-2018/31/D/NZ8/01144 funded this work at Institute of
468 Human Biology and Evolution, Faculty of Biology, Adam Mickiewicz University, 61614,
469 Poznan, Poland

470 **Datasets:** doi: 10.6084/m9.figshare.19683768

471 **References**

- 472 [1] Aiello, L. C. & Dunbar, R. I. M. 1993 Neocortex size, group size and the evolution of language.
473 *Current Anthropology* **34**, 184 - 193.
- 474 [2] Fitch, T. 2005 The evolution of language: a comparative review. *Biology and Philosophy* **20**, 193–
475 230.
- 476 [3] Pika, S., Liebal, K. & Tomasello, M. 2005 Gestural communication in subadult bonobos (*Pan*
477 *paniscus*): Repertoire and use. *American Journal of Primatology* **65**, 39-61.
- 478 [4] Seyfarth, R., Cheney, D. L. & Marler, P. 1980 Vervet monkey alarm calls: Semantic communication
479 in a free-ranging primate. *Animal Behaviour* **28**, 1070-1094.
- 480 [5] Burling, R. 1993 Primate calls, human language, and nonverbal-communication. *Current*
481 *Anthropology* **34**, 25-53.
- 482 [6] Kendon, A. 2004 *Gesture: Visible action as utterance*. Cambridge, Cambridge University Press.
- 483 [7] Pollick, A. S. & de Waal, F. B. M. 2007 Ape gestures and language evolution. *Proceedings of the*
484 *National Academy of Sciences of the United States of America* **104**, 8184-8189
- 485 [8] Cantalupo, C. & Hopkins, W. D. 2001 Asymmetric Broca's area in great apes. *Nature* **414**.
- 486 [9] de Waal, F. 2003 *Emotions inside out: 130 years after Darwin's The expression of the emotion in*
487 *man and animals*. New York, New York Academy of Science.
- 488 [10] Tomasello, M. 1999 Having intentions, understanding intentions and understanding
489 communicative intentions. In *Developing theories of intention*. (eds. P. Zelazo, D. Olson & J. Astington).
490 Cambridge, Cambridge University Press.
- 491 [11] Leavens, D. A., Russell, J. L. & Hopkins, W. D. 2005 Intentionality as measured in the persistence
492 and elaboration of communication by chimpanzees (*Pan troglodytes*). *Child Development* **76**, 291-306.
- 493 [12] Roberts, A. I., Vick, S.-J., Roberts, S. G. B. & Menzel, C. R. 2014 Chimpanzees modify intentional
494 gestures to coordinate a search for hidden food. *Nature Communications* **5**.
495 (DOI:10.1038/ncomms4088).
- 496 [13] Call, J. & Tomasello, M. 1996 The effect of humans on the cognitive development of apes. In
497 *Reaching into thought: the minds of the great apes* (eds. A. E. Russon, K. A. Bard & S. T. Parker). New
498 York, Cambridge University Press.
- 499 [14] Tomasello, M. & Call, J. 2004 The role of humans in the cognitive development of apes revisited.
500 *Animal Cognition* **7**, 213-215. (DOI:10.1007/s10071-004-0227-x).

- 501 [15] Roberts, A. I., Vick, S.-J. & Buchanan-Smith, H. 2012 Communicative intentions in wild
502 chimpanzees: Persistence and elaboration in gestural signalling. *Animal Cognition* **Published online 28**
503 **September 2012**. (DOI:10.1007/s10071-012-0563-1).
- 504 [16] McGrew, W. C. & Tutin, C. E. G. 1978 Evidence for a Social Custom in Wild Chimpanzees. *Man* **13**,
505 234-251.
- 506 [17] Leavens, D. A., Hostetter, A. B., Wesley, M. J. & Hopkins, W. D. 2004 Tactical use of unimodal and
507 bimodal communication by chimpanzees, *Pan troglodytes*. *Animal Behaviour* **67**, 467-476.
508 (DOI:10.1016/j.anbehav.2003.04.007).
- 509 [18] Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K. & Slocombe, K. E. 2013 Chimpanzee
510 alarm call production meets key criteria for intentionality. *PLoS One* **8**, e76674.
- 511 [19] Bouton, M. E. 1994 Context, ambiguity, and classical conditioning. *Current directions in*
512 *psychological science* **3**, 49-53.
- 513 [20] Whalen, P. J. 1998 Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human
514 amygdala. *Current directions in psychological science* **7**, 177-188.
- 515 [21] Gerber, A. J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J.
516 & Peterson, B. S. 2008 An affective circumplex model of neural systems subserving valence, arousal,
517 and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia* **46**, 2129-2139.
- 518 [22] Buschman, T. J. & Miller, E. K. 2007 Top-down versus bottom-up control of attention in the
519 prefrontal and posterior parietal cortices. *science* **315**, 1860-1862.
- 520 [23] Doran, D. M. & McNeilage, A. 2001 *Subspecific variation in gorilla behavior: the influence of*
521 *ecological and social factors*. New York, Cambridge University Press.
- 522 [24] Damjanovic, L., Roberts, S. G. B. & Roberts, A. 2022 Dataset for 'Language as a tool for social
523 bonding: evidence from wild chimpanzee gestural, vocal and bimodal signals'. (Figshare
524 <https://doi.org/10.6084/m9.figshare.19683768>)
- 525 [25] Roberts, A. I., Roberts, S. G. B. & Vick, S.-J. 2014 The repertoire and intentionality of gestural
526 communication in wild chimpanzees. *Animal Cognition* **17**, 317 - 336. (DOI:10.1007/s10071-013-0664-
527 5).
- 528 [26] Noudoost, B. & Moore, T. 2011 The role of neuromodulators in selective attention. *Trends in*
529 *cognitive sciences* **15**, 585-591.
- 530 [27] Westbrook, A. & Braver, T. S. 2016 Dopamine does double duty in motivating cognitive effort.
531 *Neuron* **89**, 695-710.
- 532 [28] Roberts, A. I., Murray, L. & Roberts, S. G. B. 2019 Complex sociality of wild chimpanzee can emerge
533 from laterality of manual gesture. *Human Nature*. (DOI:10.1007/s12110-019-09347-3).
- 534 [29] Fischer, J. & Price, T. 2017 Meaning, intention, and inference in primate vocal communication.
535 *Neuroscience & Biobehavioral Reviews* **82**, 22-31.
- 536 [30] Aureli, F., Preston, S. D. & de Waal, F. 1999 Heart rate responses to social interactions in free-
537 moving rhesus macaques (*Macaca mulatta*): a pilot study. *Journal of Comparative Psychology* **113**, 59.
- 538 [31] Behringer, V., Deschner, T., Möstl, E., Selzer, D. & Hohmann, G. 2012 Stress affects salivary alpha-
539 amylase activity in bonobos. *Physiology & behavior* **105**, 476-482.
- 540 [32] Kano, F., Hirata, S., Deschner, T., Behringer, V. & Call, J. 2016 Nasal temperature drop in response
541 to a playback of conspecific fights in chimpanzees: A thermo-imaging study. *Physiology & behavior*
542 **155**, 83-94.
- 543 [33] Arnsten, A. F. 2009 Stress signalling pathways that impair prefrontal cortex structure and function.
544 *Nature reviews neuroscience* **10**, 410-422.
- 545 [34] Roberts, S. G. B. & Roberts, A. I. 2016 Social brain hypothesis, vocal and gesture networks of wild
546 chimpanzees. *Frontiers in Psychology* **7**. (DOI:10.3389/fpsyg.2016.01756).
- 547 [35] Dunbar, R. I. M. 1993 Coevolution of neocortical size, group size and language in humans.
548 *Behavioral and Brain Sciences* **16**, 681-694.
- 549 [36] Davis, M. H. & Johnsrude, I. S. 2003 Hierarchical processing in spoken language comprehension.
550 *Journal of Neuroscience* **23**, 3423-3431.

551 [37] Copland, D. A., Chenery, H. J., Murdoch, B. E., Arnott, W. L. & Silburn, P. A. 2003 Dopamine
552 enhances semantic salience: semantic priming evidence from healthy individuals. *Brain and Language*
553 **1**, 103-104.

554 [38] McComb, K. & Semple, S. 2005 Coevolution of vocal communication and sociality in primates.
555 *Biology Letters* **1**, 381-385.

556 [39] Seyfarth, R. & Cheney, D. 1993 Grooming is not the only regulator of primate social interactions.
557 *Behavioral and Brain Sciences* **16**, 717-718.

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575 Table 1. Occurrence of markers of intentionality according to type of communicative
 576 behaviour

	Gesture	Vocalisation	Bimodal vocalisation with gesture	Total
Audience checking present	2269	43	274	2586
Audience checking absent	396	8	56	460
Response waiting present	1876	35	204	2115
Response waiting absent	729	18	117	864
Elaboration present	606	47	55	708
Elaboration absent	2308	31	355	2694
Approach	2103	17	250	2370
Avoidance	679	31	126	836
Initiate activity change	1539	32	218	1789
Maintain ongoing behaviour	1372	33	207	1612

577

578 Table 2. Effects of social bond of recipient towards signaller and intentionality marker
 579 (audience checking, response waiting and elaboration considered separately) including control
 580 variables (signaller and recipient sex, signaller age, relatedness) on vocal, gestural and bimodal
 581 communication.

582 a) Comparison of vocal (reference category) versus gestural signalling

Model term	Coefficient	Standard error	Significance
Signaller age [subadult]	13.174	0.814	<0.001
Relatedness [non-kin]	-12.867	1.160	<0.001
Signaller sex [female]	0.839	1.020	0.411
Recipient sex [female]	2.641	0.978	0.007
Audience checking [absent]	1.674	2.229	0.453
Response waiting [absent]	-1.611	1.801	0.372
Elaboration [absent]	4.510	0.397	<0.001
Recipient/ signaller social bond [weak]	-1.690	0.559	0.003

583

584 b) Comparison of bimodal (reference category) versus gestural signalling

Model term	Coefficient	Standard error	Significance
Signaller age [subadult]	7.754	0.597	<0.001
Relatedness [non-kin]	-0.842	0.457	0.066
Signaller sex [female]	-1.448	0.451	0.001
Recipient sex [female]	2.114	0.585	<0.001
Audience checking [absent]	1.145	0.635	0.072
Response waiting [absent]	0.144	0.226	0.525
Elaboration [absent]	0.129	0.539	0.810
Recipient/ signaller social bond [weak]	1.113	0.712	0.119

585

586 c) Comparison of bimodal (reference category) versus vocal signalling

Model term	Coefficient	Standard error	Significance
Signaller sex [female]	-0.710	1.869	0.705
Recipient sex [female]	0.131	1.697	0.939
Audience checking [absent]	-0.343	1.446	0.813
Response waiting [absent]	1.711	1.572	0.279
Elaboration [absent]	-5.061	0.773	<0.001
Recipient/ signaller social bond [weak]	3.799	0.993	<0.001

587

588 Table 3. Effects of communication type: a) vocal versus gestural, b) bimodal versus gestural
 589 and intentionality marker (audience checking, response waiting and elaboration) including
 590 control variables (signaller and recipient age, signaller and recipient sex, relatedness) on
 591 approach and avoidance response to communication. Only those events were considered,
 592 whereby the recipient was weakly bonded to the signaller.

593 a)

Model term	Coefficient	Standard error	Significance
Signaller age [subadult]	-2.516	1.367	0.067
Relatedness [non-kin]	-17.082	1.818	<0.001
Signaller sex [female]	0.233	1.347	0.863
Recipient sex [female]	1.251	1.197	0.297
Audience checking [absent]	1.961	1.648	0.236
Response waiting [absent]	-4.953	1.514	0.001
Elaboration [absent]	4.083	1.282	0.002
Communication type [vocalisation]	-0.990	0.386	0.011

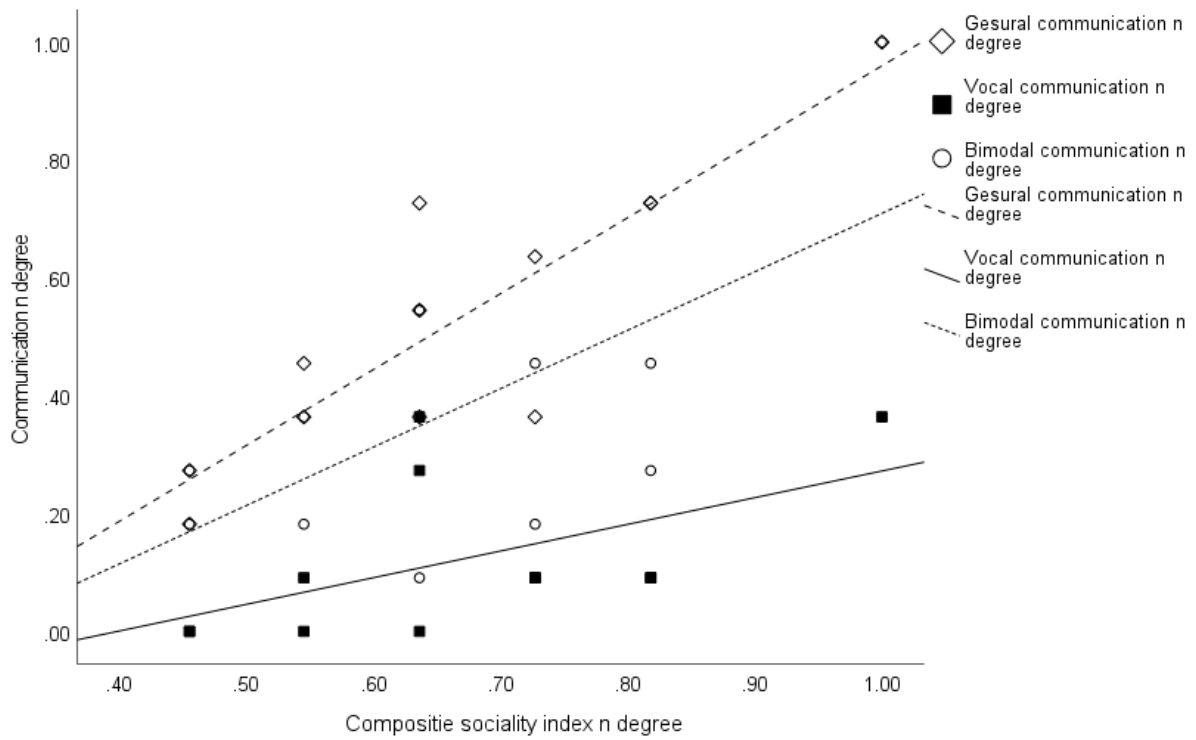
594

595 b)

Model term	Coefficient	Standard error	Significance
Signaller age [subadult]	-1.922	1.089	0.079
Relatedness [non-kin]	-1.880	1.105	0.091
Signaller sex [female]	-0.283	1.003	0.778
Recipient sex [female]	0.277	0.961	0.774
Audience checking [absent]	0.018	0.955	0.985
Response waiting [absent]	-2.174	0.847	0.011
Elaboration [absent]	2.793	1.015	0.007
Communication type [bimodal signal]	-1.332	0.647	0.041

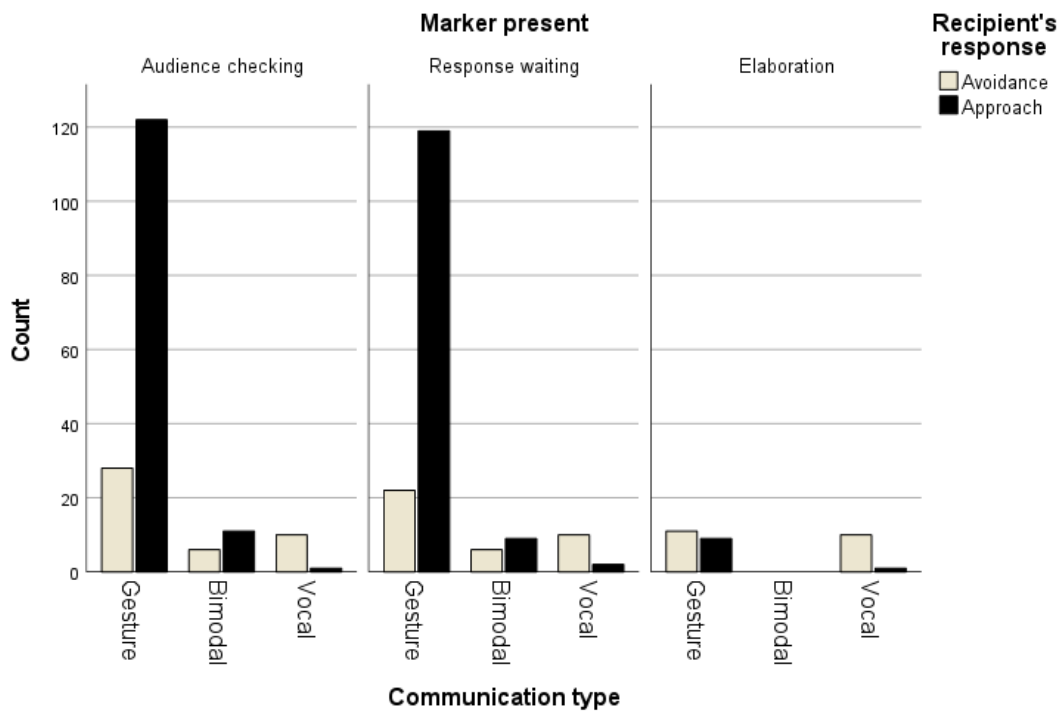
596

597 Figure 1. Relationship between communication n degree and composite sociality index n
 598 degree



599
600
601
602
603
604
605

Fig. 2 Relationship between presence of intentional signalling accompanying communication type and response type of the recipient when social bonds of the recipient towards signaller are weak



606
607
608
609

610

611

612