



**HAL**  
open science

## How's my kitty? Acoustic parameters of cat-directed speech in human-cat interactions

Charlotte de Mouzon, Caroline Gilbert, Romain Di-Stasi, Gérard Leboucher

### ► To cite this version:

Charlotte de Mouzon, Caroline Gilbert, Romain Di-Stasi, Gérard Leboucher. How's my kitty? Acoustic parameters of cat-directed speech in human-cat interactions. *Behavioural Processes*, 2022, 203, pp.104755. 10.1016/j.beproc.2022.104755 . hal-04510134

**HAL Id: hal-04510134**

**<https://hal.science/hal-04510134>**

Submitted on 26 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29

# How's my kitty?

## Acoustic parameters of cat-directed speech in human-cat interactions

Charlotte de Mouzon<sup>1</sup>, Caroline Gilbert<sup>2</sup>, Romain Di-Stasi<sup>1</sup> & Gérard Leboucher<sup>1</sup>

<sup>1</sup>UPL, Université Paris Nanterre, Laboratoire Ethologie Cognition Développement (LECD) - 200, avenue de la république - 92000 Nanterre, FRANCE. <sup>2</sup>UMR 7179, CNRS/MNHN, Ecole Nationale Vétérinaire d'Alfort (EnvA), 94700 Maisons-Alfort, FRANCE.

Corresponding author: Charlotte de Mouzon [cdemouzon@parisnanterre.fr](mailto:cdemouzon@parisnanterre.fr), +33667685130

### Keywords

Cat-directed speech; human-animal interaction; human-cat relationship; interspecific communication; Pet-directed Speech; vocal communication

## 30 **1.Introduction**

31 Domestic cats (*Felis catus*) have been sharing our environment for over ten thousand years (Hu et al.  
32 2014). Yet their relationship with humans really took a turn a few decades ago. With the major  
33 transformation of western societies mainly marked by urbanisation, human expectations of their feline  
34 companions have gradually changed. The role of cats in households is now less focused on mousing  
35 than on their companionship (Farnworth 2015). Animal companions are becoming increasingly  
36 prevalent in our lives, so much that we sometimes consider them as our “fur babies” (Greenebaum  
37 2004). As different species sharing a common ecosystem, humans and cats have developed specific  
38 behaviours to communicate with their interspecific interlocutor, enabling emitted signals to reach  
39 their targeted receiver (Schötz 2019; Turner 2021).

40 Almost thirty years ago, Levinson (1982) suggested that more empirical research needed to be  
41 directed at the study of linguistic interactions with animals. He was well advised, as it is now  
42 increasingly documented that people, in Western cultures, use a special speech register when talking  
43 to their pets, qualified as Pet-directed Speech, or PDS (Ben-Aderet et al. 2017; Benjamin & Slocombe  
44 2018; Burnham et al. 2002; Jeannin et al. 2017a, b; Lesch et al. 2019; Mitchell 2001, 2004; Ringrose  
45 2015; Xu et al. 2013). In the same way, it is rather well described that when talking to infants, human  
46 adults use a special speech register characterised, from an acoustic point of view, by an elevated  
47 fundamental frequency (pitch), exaggerated intonation contours and high affect - qualified as infant-  
48 directed Speech, or IDS (Kaplan et al. 1995). PDS shares some prosodic features of IDS, including  
49 shorter utterances, more imperative sentences, more repetitions, elevated pitch and increased pitch  
50 variation that are distinct from the typical adult-directed speech or ADS (Burnham et al. 2002; Hirsh-  
51 Pasek & Treiman 1982; Jeannin et al. 2017b; Koda 2001; Mitchell 2001; Mitchell & Edmonson 1999;  
52 Ringrose 2015; Rogers et al. 1993). Interestingly, in most of the above-mentioned studies, PDS  
53 characteristics were described using speech addressed to dogs. Much less information is available  
54 relating to cats. Sims and Chin (2002) reported that when playing with a cat, humans used very short  
55 and repetitive utterances and a large percentage of imperatives, characteristics that can also be found

56 in IDS and DDS as well (e.g., Jeannin et al. 2017b, Mitchell 2001). Nevertheless, their study did not  
57 investigate the acoustic characteristics of this speech addressed to cats. In her academic work,  
58 Acevedo (2017) reported that kitten-directed speech shows higher harmonicity than ADS. However,  
59 the differences found between ADS and cat-directed speech (CDS) remained limited, especially when  
60 speaking to an adult cat. Pitch was only marginally affected by recording conditions and speakers did  
61 not significantly modify their pitch when speaking to adult cats. Finally, in a pilot study investigating  
62 CDS in four men and four women, Schötz (2019) reported a higher mean fundamental frequency in  
63 CDS than ADS.

64 Another issue with studies on pet-directed speech is *how* the speech itself is elicited. Some  
65 experiments used photos of animals or robotic avatars (Acevedo 2017; Ben-Aderet et al. 2017; Sinatra  
66 et al. 2012), others used live human-animal interactions (Benjamin & Slocombe 2018; Burnham et  
67 al. 2002; Gergely et al. 2017, 2021; Hirsh-Pasek & Treiman 1982; Jeannin et al. 2017a, b; Koda 2001;  
68 Lesch et al. 2019; Mitchell 2001, 2004; Mitchell & Edmonson 1999; Prato-Previde et al. 2006;  
69 Ringrose 2015; Sims & Chin 2002; Sinatra et al. 2012; Xu et al. 2013). The former method is  
70 obviously simpler than the latter and allows for a better control of the experimental parameters, but  
71 it has the disadvantage of being less ecological, which could lead to less relevant results. Sinatra et  
72 al. (2012) compared human interactions with robotic entities versus live animals, finding that a real  
73 dog elicited more speech than a robot-dog. However, in their experiment, the cat was not significantly  
74 more spoken to than the robot-dog. Later on, Pepe et al. (2015) found that when talking to a live dog,  
75 participants had a higher fundamental frequency than when talking to a robotic dog. This observation  
76 suggests that findings could differ when human participants interact with a live animal vs. a robot or  
77 a picture. Indeed, the potential effect of methodological differences on characteristics of PDS was  
78 recently pointed out by Gergely et al. (2021). We have addressed this issue using both video  
79 recordings – which allow a good control of the experimental conditions but has the advantage of  
80 being more realistic than simple photographs – and live interactions with cats, to elicit speech directed  
81 toward cats.

82 In this context, the aim of the present study was to investigate parameters of cat-directed speech,  
83 focusing on acoustic characteristics. Our primary hypothesis was that CDS displays a higher pitch  
84 than ADS. Secondly, based on previous findings that women used the baby talk speech register  
85 more than men when talking to dogs (Lesch et al. 2019, Mitchell 2001, 2004; Prato-Previde et al.  
86 2006), we predicted that there would be differences between men and women's speech addressed to  
87 cats. We tested these hypotheses using video recordings in standardised laboratory settings  
88 (experiment 1) and in more ecological conditions through direct human-cat interactions (experiment  
89 2), to investigate whether experimental conditions would affect the results.

90

## 91 **2. Methods**

### 92 **2.1. Subjects**

93 All participants were naïve to the experiment. In order for them to be as spontaneous as possible, the  
94 main objectives of the study were only briefly described in the consent form; participants received  
95 more specific information about the aims of the study at the end of the experiment.

96 For the video condition (experiment 1), thirty-two participants were recruited, on a voluntary basis,  
97 among students of Paris-Nanterre University (Sports, Sciences and Psychology departments). In order  
98 to test for gender specificity of the effects, the cohort included 16 women and 16 men of similar age  
99 ( $20.5 \pm 0.2$  vs.  $21.4 \pm 0.6$  years respectively, Mann-Whitney  $U = 99.00$ ,  $p = 0.23$ ).

100 For direct human-cat interaction (experiment 2), we studied participants interacting with their own  
101 cat. Women were recruited on a voluntary basis among students of Ecole nationale vétérinaire  
102 d'Alfort, (EnvA). Twenty-three students owning a cat showed interest in participating (twenty-one  
103 women and two men). Applicants whose cats were aged under 6 months, presented significant health  
104 or hearing issues, aggressiveness, or exaggerated inhibition were excluded from the study. This was  
105 determined by declarative statement of their owners. Because of the gender disparity in the  
106 participants recruited at EnvA, men were excluded from the analysis for the sake of homogeneity.

107 The final number of participants was fifteen women, aged  $22.4 \pm 0.16$  years and their cats aged 8  
108 months to 2.5 years, (mean  $1.5 \pm 0.13$  years). Seven cats were spayed females, seven were neutered  
109 males and one was an intact male.

#### 110 *Ethics approval*

111 All procedures in the present study were in accordance with the ethical standards of the institutional  
112 committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical  
113 standards. All applicable international, national and/or institutional guidelines for the care and use of  
114 animals were followed and all procedures performed in studies involving cats were in accordance  
115 with the ethical standards of the institution at which the study was conducted. The following protocol  
116 was approved by the Ethics Committee for Clinical Research (Comité d’Ethique en Recherche  
117 Clinique, ComERC) of EnvA, Saisine n°2018-10-24.

#### 118 *Consent to participate*

119 Informed consent was obtained from all individual participants included in the study. The main  
120 objectives of the study were described in the consent form.

121

## 122 **2.2. Experimental settings and data collection**

123 For both experiments, voices of human participants were recorded with a lapel microphone (Rode  
124 SmartLav+®) connected to a digital recorder (MARANTZ PMD620®). Stimuli were recorded as  
125 uncompressed wav Audio files at a sampling rate of 48 kHz and a resolution of 16 bits. The sensitivity  
126 was adjusted prior to recording to prevent clipping.

### 127 **2.2.1 Experiment 1:**

128 The investigation took place at the LECD (Laboratoire Ethologie Cognition Développement), Paris  
129 Nanterre University, in a quiet room. Participants were presented a 20 seconds video of a cat not  
130 paying attention to the camera, on a laptop computer (HP 15 Notebook®). In order to avoid pseudo-  
131 replication bias (Kroodsma 1989) we prepared a set of 16 cat-videos, so each female and male

132 participant would see a different video from other participants of the same gender. The experimenter  
133 started the video, introduced the cat by its name and asked the participant to remind her what was the  
134 cat's name (ADS register) and to call the cat by its name in order to get its attention (assumed CDS).  
135 All cats on the videos were active, i.e., watching tv, playing with a toy, looking at the window, passing  
136 by the camera, etc., to motivate the participant to catch their attention. ADS and CDS production  
137 were counterbalanced randomly (experimenter elicited either ADS or CDS first).

### 138 ***2.2.2 Experiment 2:***

139 The investigation took place at the Ecole nationale vétérinaire d'Alfort, (EnvA), in a room previously  
140 used for acoustic experiments with dogs (Jeannin et al. 2017a, b). The 24 m<sup>2</sup> room was adapted to  
141 become "cat friendly", with numerous hiding spots, height-access, a large litter box, various  
142 scratchers and games that best fit feline behavioural needs (Ellis 2009). Most equipment was supplied  
143 by Martin Sellier®, Valenciennes, France. A few days before the experiment, human and cat  
144 participants were allowed to freely spend two 1-hour visits in the room, in order to familiarise them  
145 with the setting conditions and experimenters. In order for the participants to feel as comfortable as  
146 possible, random discussions, free interactions and playing with cats were encouraged. Two female  
147 experimenters, Experimenter 1 and Experimenter 2, were involved in the study. Experimenters and  
148 participants spoke French through the whole experiment.

#### 149 ***2.2.2.1 Preliminary phase***

150 Experimenters greeted participants, and cats were let out of their carrier so they could explore the  
151 room freely and feel comfortable in the experimental setting. Experimenter 1 explained the different  
152 steps of the study, provided the microphone and started audio recording. Therefore, participants were  
153 recorded when talking to a human adult (Experimenter 1), or to their cat at all times, whether the  
154 occurrences were spontaneous or not. A series of questions were asked, regarding the cat's  
155 environment, habits and behaviour. Then, unscripted free-flowing conversation took place so that  
156 human participants would feel comfortable with experimenters. The total duration of this preliminary  
157 phase was about 20 min.

158 2.2.2.2. *Experimental phases*

159 Three conditions of interaction were established, so speech from the owner to the cat in different  
160 types of interactions could be audio-recorded. The owner was asked to interact with her cat in the  
161 same manner as she would at home. The order of testing was (1) play and treat, (2) separation, (3)  
162 reunion.

163 2.2.2.2.1. Play and treat condition:

164 In a counterbalanced order, participants were asked to play with their cats and to give them a treat:  
165 half of the participants played first, the other half gave a treat first. While reproducing these everyday  
166 life conditions, experimenter 1 explicitly asked participants to say to their cat: “tu veux jouer ?” [*ty*  
167 *vø ʒwe*] (do you want to play?) or “tu veux manger ?” [*ty vø mɑ̃.ʒe*] (do you want a treat?); different  
168 toys and treats were available in the room; owners could also use their own if preferred. Play and treat  
169 condition lasted 10 minutes.

170 2.2.2.2.2 Separation condition:

171 The owner and experimenter 1 left the room while experimenter 2 stayed with the cat. Before leaving,  
172 the owner was asked to say “à tout à l’heure” [*a tu.t\_a l\_œʁ*] (see you later) to her cat, in the way  
173 she usually does. While outside the experimental room, experimenter 1 asked the owner to say to her  
174 the four standardised sentences, as if the experimenter was a friend or family member: “do you want  
175 to play?”, “do you want a treat?”, “how are you?” and “see you later”. Separation condition lasted 5  
176 minutes.

177 2.2.2.2.3. Reunion condition:

178 The owner and experimenter returned to the room and the owner was free to greet and interact with  
179 her cat as she would normally do. Experimenter 1 specifically asked the owner to say to her cat  
180 “comment ça va ?” [*kɔ.mɑ̃ sa va*] (how are you?). Reunion condition lasted 5 minutes.

181

182



### 183 2.3. Acoustic analysis

184 Audacity® recording and editing software 2.2.0 were used to isolate utterances addressed either to  
185 cats (assumed CDS) or to human adults (ADS). For the first experiment two fragments, the cat’s name  
186 in CDS and the cat’s name in ADS, were isolated for each participant. For the second experiment,  
187 eight fragments, four CDS and four ADS, were isolated for each woman, that is one fragment in ADS  
188 and one fragment in CDS for each type of sentence. In order to compare the four standardised  
189 sentences with random spontaneous utterances addressed to their cats, we also extracted utterances  
190 of spontaneous CDS fragments and aggregated them into a unique audio sequence, using Audacity®.  
191 The spontaneous CDS utterances were randomly extracted from 15-minute fragments of the  
192 recordings, i.e., we randomly selected 15-minute audio extracts for each participant of experiment 2  
193 and selected all utterances addressed by the participant to her cat within these 15 minutes. The four  
194 standardised sentences uttered in CDS were also chained together for the purpose of this comparison.

195 Acoustic analyses were performed using PRAAT software 6.0.37 (Boersma 2001; Boersma &  
196 Weenink 2018), using the standard settings. Based on Jeannin et al.’s (2017b) findings in dogs, three  
197 parameters were isolated: (a) mean F0, the average fundamental frequency over the duration of the  
198 signal, which represents the pitch; (b) F0CV: the coefficient of variation of F0 over the duration of  
199 the signal, estimated as the standard deviation of F0 (F0SD) divided by mean F0. F0CV is a measure  
200 of the intonation and represents the pitch modulation; (c) IntCV: the coefficient of variation of the  
201 intensity contour, which represents the modulation of the energy in the sound. Mean F0 and F0\_SD  
202 were measured using the “Voice report” function in PRAAT. Pitch settings were 75-600 Hz and  
203 fragments containing the considered ADS and CDS utterances were selected. F0CV was calculated  
204 as F0SD/mean F0. IntCV was calculated as IntSD/mean Int. Mean intensity (mean Int) was measured  
205 using the “Get intensity” function, standard deviation of intensity (IntSD) was measured using the  
206 query “Get standard deviation”.

207

208

## 209 **2.4. Human evaluation of women's utterances**

210 In order to confirm that audio recordings of CDS collected through direct human-cat interactions  
211 would also be identified by naïve human raters as different from ADS, we chained together a  
212 randomised sample of the 8 standardised sentences, i.e., 4 ADS and 4 CDS, pronounced by all fifteen  
213 women recorded at EnvA. Fragments were aggregated together using Audacity® recording and  
214 editing software 2.2.0., with randomised CDS and ADS fragments alternated with 2-second silences,  
215 forming a 4-minute audio sequence. 15 independent raters – 3 men and 12 women (age  $31.9 \pm 3$   
216 years,) were presented the same 71 chained fragments but in a different order. They were asked to  
217 rate fragments as ADS or CDS, hearing them only once.

218

## 219 **2.5. Statistical analysis**

220 In the first experiment, because all data did not follow statistical normality (Shapiro-Wilk test for  
221 normality  $p < 0.05$ ), we used non-parametric Wilcoxon signed rank test to compare acoustical  
222 characteristics of ADS and CDS utterances (Siegel & Castellan 1988). These analyses were made  
223 using SigmaPlot® 13.0 software. Acoustic characteristics of men and women's speech being clearly  
224 different (Torre & Barlow 2009), we analysed them separately. Two-tailed tests were used  
225 throughout.

226 In the second experiment, the effect of speech register (ADS vs. CDS) and phrase types (“do you  
227 want to play?”, “do you want a treat?”, “how are you?” and “see you later”) were analysed with linear  
228 mixed-effects models (LMM) using the *lme* function of the R package *nlme* (Pinheiro et al. 2019).  
229 Individual identity was included as a random factor. P-values were calculated with a permutation test  
230 (Monte Carlo sampling with 1000 permutations), using the *PermTest* function of the R package  
231 *pgirmess* (Giraudoux 2018). This statistical procedure can be performed even when normality is not  
232 present. If the results of the first analyses were significant, post-hoc analyses were performed using  
233 the *PermTest* function. For each post-hoc comparison, p-values were compared to alpha risk adjusted

234 according to the number of pairwise comparisons with the Bonferonni sequential correction (Holm-  
235 Bonferroni method). These analyses were done with R, version 3.6.0 (R Core Team 2019).  
236 Comparison of standardised CDS sentences versus spontaneously uttered CDS was made using a  
237 Wilcoxon signed rank test (Siegel & Castellan 1988). This same test was performed for human  
238 evaluation of women's utterances, to evaluate the difference between the number of correct answers  
239 and wrong answers given by naïve human raters. These analyses were made using SigmaPlot® 13.0  
240 software. We used Matched Pair Rank Biserial Correlation  $r$  to evaluate the effect size after non-  
241 parametric analyses (Tomczak & Tomczak, 2014), and Cohen's  $d$  after parametric analyses (Goulet-  
242 Pelletier & Cousineau, 2018). These analyses were performed using JASP 0.16.2 software.

243

## 244 **3. Results**

### 245 *3.1. Experiment 1:*

246 Mean fundamental frequency (mean F0), variation of frequency (F0CV) and variation of intensity  
247 contour (IntCV) were compared for ADS and CDS condition of utterances elicited in a standardised  
248 laboratory setting using video recordings. Detailed results are presented in Table 1 and data split  
249 according to gender of human participants are presented in Figure 1. Women had a significantly  
250 higher mean F0 in CDS than in ADS condition. Women also had a higher IntCV in CDS than in ADS  
251 condition. In contrast, no significant difference was found between CDS and ADS for F0CV. Men  
252 had a higher mean F0 in CDS than in ADS condition. However, there were no difference between  
253 CDS and ADS for IntCV or for F0CV.

### 254 *3.2. Experiment 2:*

255 We analysed the same acoustic parameters of ADS and CDS register, i.e. mean F0, F0CV and IntCV,  
256 of participants interacting with another adult human or with their own cat, all being present in the  
257 same room. Results of the LMM are presented in Table 2. For mean F0, there was a significant effect  
258 of speech register ( $p < 0.001$ , see Figure 2): mean F0 was higher in the CDS than ADS condition

259 (308.3  $\pm$  6.38 Hz vs. 230.52  $\pm$  4.29 Hz respectively), (Cohen's  $d = 1.656$ ). In contrast there was no  
260 significant effect of the type of sentence nor a significant interaction between the two factors. The  
261 analysis also indicates that F0CV was significantly influenced by the type of sentence ( $p = 0.004$ ) but  
262 there was no significant effect of speech register nor a significant interaction between the two factors;  
263 post-hoc analyses revealed that participants had a higher F0CV when uttering the sentence "do you  
264 want to play?" than for each of the other sentences i.e. "how are you" ( $p < 0.001$ ), "see you later" ( $p$   
265  $= 0.001$ ) and "do you want a treat?" ( $p = 0.003$ ). In addition, the sentence "do you want a treat?"  
266 presented a higher F0CV than "how are you?" ( $p = 0.002$ ). For IntCV, there was no significant effect  
267 of the speech register nor of the type of sentence; however, the interaction between these two factors  
268 was significant ( $p = 0.043$ ). Post-hoc analyses failed to reveal any significant difference (at least,  $p >$   
269  $0.1$ ).

270 In order to confirm that the four standardised CDS sentences were equivalent to regular speech  
271 addressed to cats, we compared 15 aggregated fragments of standardised sentences (mean duration =  
272  $3.09 \pm 0.19$  seconds) to 15 aggregated fragments of spontaneous utterances addressed by the same  
273 human participants to their cats (mean duration =  $5.08 \pm 0.29$  seconds). There were no significant  
274 differences between spontaneous CDS and standardised sentences for Mean F0, F0CV and IntCV.  
275 Following results are given as median values. Mean F0: 302.45 vs. 312.00; Wilcoxon  $z = 0.17$ ,  $p =$   
276  $0.89$ , Matched pair rank biserial correlation,  $r = 0.05$ ; F0CV: 0.23 vs. 0.21;  $z = -0.11$ ,  $p = 0.93$ ,  $r = -$   
277  $0.033$ ); IntCV: 1.33 vs. 1.28;  $z = -0.80$ ,  $p = 0.45$ ,  $r = -0.23$ . Acoustic evaluation of ADS and CDS by  
278 independent human raters revealed that the speech register was correctly identified in  $61.7 \pm 0.7$  cases  
279 out of 71, i.e. 86.9 %. The number of correct answers was significantly higher than the number of  
280 wrong answers (Wilcoxon signed rank test,  $z = -3.417$ ,  $p < 0.001$ ,  $r = 1$ ) indicating that raters did not  
281 answer randomly.

282 Overall, data collected from both experiments confirm our primary hypothesis that humans use a  
283 specific vocal register when addressing cats, mainly characterised by a higher pitch in CDS than in  
284 ADS. For all data, the effect size analyses are in line with the results of the mean comparisons, with

285 large values when the means differ significantly. Intergroup differences therefore appear to make  
286 sense.

287

## 288 **4. Discussion**

289 We report here that humans address cats using a particular speech with a higher pitch than they use  
290 when talking to adult humans. Under two different experimental conditions, in two different  
291 populations, our findings reinforce Schötz (2019) reports from her pilot study. In their work with a  
292 small cohort of dogs and cats, Burnham et al. (2002) used the broad term of pet-directed speech and  
293 did not discriminate speech addressed to cats and dogs. As recent studies now specifically mention  
294 dog-directed speech, or DDS (Ben-Aderet et al. 2017; Benjamin & Slocombe 2018 Gergely et al.  
295 2017; Gergely et al. 2021; Ringrose 2015; Xu et al. 2013), we chose, as did Schötz (2019), to refer to  
296 the speech addressed to cats as cat-directed speech (CDS).

297 In the video condition, we observed a higher pitch (mean F0) in the CDS than in the ADS condition,  
298 regardless of the participant's gender. In a separate cohort allowing direct owner-cat interactions, we  
299 confirmed that women interacting with their own cat had a higher pitch in the CDS than in the ADS  
300 condition. To this regard, video recordings and direct interactions generate similar results. These high  
301 speech features are in line with the acoustical characteristics previously reported for speech addressed  
302 to dogs, with women (Ben-Aderet et al. 2017; Gergely et al. 2017; Jeannin et al. 2017a) as well as  
303 men (Gergely et al. 2017) displaying a higher pitch in DDS than in ADS conditions.

304 In addition to the high speech feature, the variation of intensity contour (IntCV) was higher in the  
305 CDS than in the ADS condition when women talked to an unknown cat (video recording). This was  
306 not observed in men, suggesting that female participants displayed a more explicit CDS than male  
307 participants. This is in line with previous findings reporting differences between men and women  
308 addressing dogs (Mitchell 2001, 2004; Prato-Previde et al. 2006). For example, Mitchell (2001)  
309 showed that women used a baby talk speech register more frequently than did men. Also, Prato-

310 Previde et al. (2016) reported that women's utterances resembled more closely IDS than did men's  
311 utterances, that women talked more than men, and that women had a shorter latency in starting talking  
312 to dogs. This slight difference between men and women might relate to men feeling more constrained  
313 in speech than women and thus repressing their expression. In experiment 2, through direct owner-  
314 cat interaction, female participants did not increase the modulation of their voice intensity when  
315 talking to their own cat compared to an adult human. This finding is consistent with data collected in  
316 dogs under similar experimental conditions (Jeannin et al. 2017a). This slight discrepancy between  
317 experiment 1 and experiment 2 is in line with a previous study, in which participants used more DDS  
318 when interacting with an unfamiliar dog than with a familiar dog (Mitchell 2004).

319 The variation of fundamental frequency (F0CV) was not modified by speech register, in women nor  
320 in men, under both experimental settings, suggesting that adults do not increase the modulation of  
321 their pitch when talking to cats. This finding appears to be specific to cats, as it differs from results  
322 reported in dogs, under similar experimental conditions used for experiment 2 (Jeannin et al. 2017a).  
323 Our results might be considered in the light of Xu et al. (2013)'s findings, who reported that the  
324 degree of vowel hyperarticulation increased from ADS and DDS to parrot-directed speech, then to  
325 IDS. Xu and colleagues hypothesised that the degree of hyperarticulation might be related to the  
326 audience's actual or expected linguistic competence. Hyperarticulation and pitch variation both play  
327 a role in speech expressivity (Audibert et al. 2005; Beller et al. 2008). Therefore, if the production of  
328 vowel hyperarticulation depends on humans' perception of their interlocutor's ability to understand  
329 their communicative intent, we could hypothesize that a greater pitch variation in DDS than CDS  
330 would reflect that humans expect more understanding or feedback from dogs than cats. To test this  
331 hypothesis, further experiments should compare CDS to DDS in the same study cohort. Interestingly,  
332 pitch modulation was significantly influenced by the type of sentence pronounced by the participants.  
333 Post-hoc analyses revealed that when uttering the phrase "do you want to play?" participants had a  
334 higher pitch modulation than for each of the other sentences. In addition, the sentence "do you want  
335 a treat?" presented a higher pitch modulation than "how are you?". In line with the above hypothesis,

336 one could postulate that an emitter, regardless of the receiver's species, might want to increase  
337 expressivity when asking a question expecting a reaction rather than when simply expressing  
338 greetings.

339 Finally, when comparing audio extracts recorded in the second experiment, we did not find any  
340 difference between spontaneous speech and standardised sentences for mean F0, F0CV and IntCV.  
341 This suggests that our findings, obtained with standardised sentences, are consistent with what occurs  
342 in everyday spontaneous interactions. Furthermore, we found that naïve human raters correctly  
343 identified the speech register of unknown speakers. This observation tends to reinforce the credibility  
344 of audio extracts analysed in the present experiment.

345 The role of CDS has not been investigated yet and this particular kind of speech may have emerged  
346 through various routes during the domestication process. The development of closer human-cat  
347 relationships within the last decades (Turner 2017) may be one reason for the growing use of this  
348 speech register. Indeed, it is thought that infant directed speech (IDS) increases social bonding  
349 between infant and caregiver (Kaplan et al. 1995). Benjamin & Slocombe (2018) suggested that dog  
350 directed speech (DDS) may fulfil a dual function of strengthening the affiliative bond and improving  
351 the listener's attention. This hypothesis is consistent with earlier findings by Jeannin and co-workers  
352 (2017a), who reported that DDS draws dogs' attention more efficiently than ADS. Earlier on, Mitchell  
353 (2001) also noted that "baby talk" to dogs was used for attention-getting and expressed friendliness  
354 and affection. Additionally, Schötz (2019) pointed out that elevated pitch is usually associated with  
355 "small" meanings, including vulnerability, uncertainty, and friendliness, which also relates to  
356 encouraging a closer interaction between partners using a higher pitch. Indeed, Schötz (2019) reported  
357 a higher mean fundamental frequency in interspecific than intraspecific utterances for both humans  
358 and cats. Taken together, there is a growing agreement for relating speech directed to animal  
359 companions, to attention-getting and attachment mechanisms. Topál and co-workers (1998) showed  
360 that adult dogs could be categorised along the secure-insecure attachment dimensions of Ainsworth's  
361 original test (Ainsworth & Bell 1970). Under similar experimental conditions, Vitale et al. (2019)

362 reported that cats also display secure, ambivalent, and avoidant attachment styles toward human  
363 caregivers, which appear to be relatively stable and present in adulthood. These latest data support  
364 the hypothesis that, just as their canine counterparts, domestic cats display an ability to form strong  
365 bonding with human caregivers. The fact that, in return, humans might use a specific vocal register  
366 to strengthen this bond, seems only to be expected. Further research could investigate a potential link  
367 between the use of CDS by humans and their attachment to their cats. To this end, it would be  
368 interesting to combine the use of thorough questioning tools, such as the cat-owner relationship scale  
369 (CORS, Howell et al. 2017), to biologically relevant acoustic measures.

370



371 **Acknowledgements**

372 We would like to thank Laetitia Matray for helping to set up the cat friendly room at the veterinary  
373 school of Maisons-Alfort (EnvA), Chloe Tavernier and Benoit Janicot for their help with the data  
374 collection, members of the LECD for their participation at different stages of the study and Sylvie  
375 Hauguel for thorough proofreading. Many thanks to all human and cat participants.

376

377 **Funding**

378 This work was supported by the French Ministry of Research and Higher Education subsidies  
379 granted to MARTIN SELLIER®

380

381 **Research data**

382 The data sets generated and analysed during the current study are available from the corresponding  
383 author on reasonable request.

384

385 **Declarations of interest:** none

386

387

388 **References**

- 389 Acevedo, D. (2017). The Association of Feline Behavior to Acoustical Features of Kitten Directed  
390 Speech. *CUNY Academic Works*. [http://academicworks.cuny.edu/hc\\_sas\\_etds/184](http://academicworks.cuny.edu/hc_sas_etds/184)
- 391 Ainsworth, M.D.S., & Bell, S.M. (1970). Attachment, Exploration, and Separation: Illustrated by the  
392 Behavior of One-Year-Olds in a Strange Situation. *Child Development* 41, 49–67.  
393 <https://doi.org/10.2307/1127388>
- 394 Audibert, N., Aubergé, V., & Rilliard, A. (2005). The prosodic dimensions of emotion in speech: the  
395 relative weights of parameters. In *Ninth European Conference on Speech Communication and*  
396 *Technology*. <http://dx.doi.org/10.21437/Interspeech.2005-332>
- 397 Barton, K., & Barton, M.K. (2019). Package ‘MuMIn’. R package version 1.43.15. [https://CRAN.R-](https://CRAN.R-project.org/package=MuMIn)  
398 [project.org/package=MuMIn](https://CRAN.R-project.org/package=MuMIn)
- 399 Beller, G., Obin, N., & Rodet, X. (2008). Articulation degree as a prosodic dimension of expressive  
400 speech. In *in Fourth International Conference on Speech Prosody*.
- 401 Ben-Aderet, T., Gallego-Abenza, M., Reby, D., & Mathevon, N. (2017). Dog-directed speech: Why  
402 do we use it and do dogs pay attention to it? *Proceedings of the Royal Society* 284, 20162429.  
403 <https://doi.org/10.1098/rspb.2016.2429>
- 404 Benjamin, A., & Slocombe, K. (2018). ‘Who’s a good boy?!’ Dogs prefer naturalistic dog-directed  
405 speech. *Animal Cognition* 21, 353–364. <https://doi.org/10.1007/s10071-018-1172-4>
- 406 Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What’s new, pussycat? On talking to  
407 babies and animals. *Science* 296,1435. <https://doi.org/10.1126/science.1069587>
- 408 Ellis, S. L. (2009). Environmental enrichment: practical strategies for improving feline welfare.  
409 *Journal of Feline Medicine and Surgery*, 11(11), 901-912.  
410 <https://doi.org/10.1016/j.jfms.2009.09.011>

- 411 Farnworth, M.J. (2015). Cats have many lives: Applying behaviour to explore the cat-human  
412 relationship. *Applied Animal Behaviour Science* 173,1–2.  
413 <https://doi.org/10.1016/j.applanim.2015.08.029>
- 414 Gergely, A., Faragó, T., Galambos, Á., & Topál, J. (2017). Differential effects of speech situations  
415 on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Scientific*  
416 *Reports* 7, 13739. <https://doi.org/10.1038/s41598-017-13883-2>
- 417 Gergely, A., Tóth, K., Faragó, T., & Topál, J. (2021). Is it all about the pitch? Acoustic determinants  
418 of dog-directed speech preference in domestic dogs, *Canis familiaris*. *Animal Behaviour* 176,  
419 167–174. <https://doi.org/10.1016/j.anbehav.2021.04.008>
- 420 Giraudoux, P., Antonietti, J.P., Beale, C., Pleydell, D., & Treglia, M. (2018). pgirmess: Spatial  
421 analysis and data mining for field ecologists. R Package version 1.6.9. [https://CRAN.R-](https://CRAN.R-project.org/package=pgirmess)  
422 [project.org/package=pgirmess](https://CRAN.R-project.org/package=pgirmess)
- 423 Goulet-Pelletier, J. C. & Cousineau, D. (2018). A review of effect sizes and their confidence intervals,  
424 Part I: The Cohen's d family. *The Quantitative Methods for Psychology*, 14(4), 242-265.
- 425 Greenebaum, J. (2004). It's a Dog's Life: Elevating Status from Pet to « Fur Baby » at Yappy Hour.  
426 *Society & Animals* 12, 117–135. <https://doi.org/10.1163/1568530041446544>
- 427 Hirsh-Pasek, K., & Treiman, R. (1982). Doggerel: Motherese in a new context. *Journal of Child*  
428 *Language* 9: 229–237. <https://doi.org/10.1017/S0305000900003731>
- 429 Howell, T. J., Bowen, J., Fatjó, J., Calvo, P., Holloway, A., & Bennett, P. C. (2017). Development of  
430 the cat-owner relationship scale (CORS). *Behavioural Processes*, 141, 305-315.  
431 <https://doi.org/10.1016/j.beproc.2017.02.024>
- 432 Hu, Y., Hu, S., Wang, W., Wu, X., Marshall, F.B., Chen, X., Hou, L., & Wang, C. (2014). Earliest  
433 evidence for commensal processes of cat domestication. *Proceedings of the National Academy*  
434 *of Sciences* 111, 116–120. <https://doi.org/10.1073/pnas.1311439110>

- 435 Jeannin, S., Gilbert, C., Amy, M., & Leboucher, G. (2017a). Pet-directed speech draws adult dogs'  
436 attention more efficiently than Adult-directed speech. *Scientific Reports* 7, 4980.  
437 <https://doi.org/10.1038/s41598-017-04671-z>
- 438 Jeannin, S., Gilbert, C., & Leboucher, G. (2017b). Effect of interaction type on the characteristics of  
439 pet-directed speech in female dog owners. *Animal Cognition* 20, 499–509.  
440 <https://doi.org/10.1007/s10071-017-1077-7>
- 441 JASP Team (2022). JASP (Version 0.16.2)[Computer software].
- 442 Kaplan, P.S., Goldstein, M.H., Huckleby, E.R., Owren, M.J., & Cooper, R.P. (1995). Dishabituation  
443 of visual attention by infant- versus adult-directed speech: Effects of frequency modulation and  
444 spectral composition. *Infant Behavior & Development* 18, 209–223.  
445 [https://doi.org/10.1016/0163-6383\(95\)90050-0](https://doi.org/10.1016/0163-6383(95)90050-0)
- 446 Koda, N. (2001). Anthropomorphism in Japanese women's status terms used in talk to potential guide  
447 dogs. *Anthrozoös* 14(2), 109–111. <https://doi.org/10.2752/089279301786999490>
- 448 Kroodsma, D. E. (1989). Suggested experimental designs for song playbacks. *Animal Behaviour* 37,  
449 600–609. [https://doi.org/10.1016/0003-3472\(89\)90039-0](https://doi.org/10.1016/0003-3472(89)90039-0)
- 450 Lesch, R., Kotrschal, K., Schöberl, I., Beetz, A., Solomon, J., & Fitch, W.T. (2019). Talking to Dogs:  
451 Companion Animal-Directed Speech in a Stress Test. *Animals*, 9(7), 417.  
452 <https://doi.org/10.3390/ani9070417>
- 453 Levinson, B. M. (1982). The Future of Research into Relationships Between People and Their Animal  
454 Companions. *International Journal for the Study of Animal Problems* 3(4), 283–294.
- 455 Mitchell, R.W. (2001). Americans' Talk to Dogs: Similarities and Differences With Talk to Infants.  
456 *Research on Language & Social Interaction*, 34(2), 183–210.  
457 [https://doi.org/10.1207/S15327973RLSI34-2\\_2](https://doi.org/10.1207/S15327973RLSI34-2_2)
- 458 Mitchell, R.W. (2004). Controlling the dog, pretending to have a conversation, or just being friendly?  
459 Influences of sex and familiarity on Americans' talk to dogs during play. *Interaction Studies*.

460 *Social Behaviour and Communication in Biological and Artificial Systems*, 5(1), 99–129.  
461 <https://doi.org/10.1075/is.5.1.06mit>

462 Mitchell, R. W., & Edmonson, E. (1999). Functions of repetitive talk to dogs during play: Control,  
463 conversation, or planning? *Society and Animals* 7(1), 55–81.  
464 <https://doi.org/10.1163/156853099X00167>

465 Pepe, A. A., Ellis, L. U., Sims, V. K., & Chin, M. G. (2008). Go, Dog, Go: Maze Training AIBO vs.  
466 a Live Dog, An Exploratory Study. *Anthrozoös*, 21(1), 71–83.  
467 <https://doi.org/10.2752/089279308X274074>

468 Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2019). R Core Team. nlme: linear and nonlinear  
469 mixed effects models. R package version 3.1-139. <https://CRAN.R-project.org/package=nlme>

470 Prato-Previde, E., Fallani, G., & Valsecchi, P. (2006). Gender Differences in Owners Interacting with  
471 Pet Dogs: An Observational Study. *Ethology* 112, 64–73. [https://doi.org/10.1111/j.1439-](https://doi.org/10.1111/j.1439-0310.2006.01123.x)  
472 [0310.2006.01123.x](https://doi.org/10.1111/j.1439-0310.2006.01123.x)

473 R Core Team. (2019). R: A language and environment for statistical computing. *R Foundation for*  
474 *Statistical Computing, Vienna, Austria*. <https://www.R-project.org/>

475 Ringrose, C.C. (2015). Pitch Change in Dog-Directed Speech. *Lifespans and Styles* 1, 28–35.  
476 <https://doi.org/10.2218/ls.v1i0.2015.1181>

477 Rogers, J., Hart, L. A., & Boltz, R. P. (1993). The role of pet dogs in casual conversations of elderly  
478 adults. *The Journal of Social Psychology* 133(3), 265–277.  
479 <https://doi.org/10.1080/00224545.1993.9712145>

480 Schötz, S. (2019). Paralinguistic information and biological codes in intra- and interspecific vocal  
481 communication: A pilot study of humans and domestic cats. *Proceedings from FONETIK 2019*,  
482 Stockholm, pp. 67-72. <https://doi.org/10.5281/zenodo.3246003>

483 Sims, V.K., & Chin, M.G. (2002). Responsiveness and perceived intelligence as predictors of speech  
484 addressed to cats, *Anthrozoös* 15:2, 166-177. <http://dx.doi.org/10.2752/089279302786992667>

- 485 Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd ed.)  
486 *New York: McGraw-Hill.*
- 487 Sinatra, A. M., Sims, V. K., Chin, M. G., & Lum, H. C. (2012). If it looks like a dog: The effect of  
488 physical appearance on human interaction with robots and animals. *Interaction Studies. Social*  
489 *Behaviour and Communication in Biological and Artificial Systems* 13(2), 235–262.  
490 <https://doi.org/10.1075/is.13.2.04>
- 491 Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview  
492 of some recommended measures of effect size. *Trends in Sport Sciences*, 1(21), 19-25.
- 493 Topál, J., Miklósi, Á., Csányi, V., & Dóka, A. (1998). Attachment Behavior in Dogs (*Canis*  
494 *familiaris*): A New Application of Ainsworth's (1969) Strange Situation Test. *Journal of*  
495 *Comparative Psychology* 112(3), 219–229. <https://doi.org/10.1037/0735-7036.112.3.219>
- 496 Torre, P., & Barlow, J.A. (2009). Age-related changes in acoustic characteristics of adult speech.  
497 *Journal of Communication Disorders* 42, 324–333.  
498 <https://doi.org/10.1016/j.jcomdis.2009.03.001>
- 499 Turner, D.C. (2017). A review of over three decades of research on cat-human and human-cat  
500 interactions and relationships. *Behavioural Processes* 141, 297–304.  
501 <https://doi.org/10.1016/j.beproc.2017.01.008>
- 502 Turner, D.C. (2021). The Mechanics of Social Interactions Between Cats and Their Owners.  
503 *Frontiers in Veterinary Science* 8: 650143. <https://doi.org/10.3389/fvets.2021.650143>
- 504 Vitale, K.R., Behnke, A.C., & Udell, M.A.R. (2019). Attachment bonds between domestic cats and  
505 humans. *Current Biology* 29, 864–R865. <https://doi.org/10.1016/j.cub.2019.08.036>
- 506 Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013). Vowel Hyperarticulation in Parrot-  
507 Dog- and Infant- Directed Speech. *Anthrozoös*, 26(3), 373–380.  
508 <https://doi.org/10.2752/175303713X13697429463592>
- 509

510 **Table 1** Effect of speech register (ADS vs. CDS) on mean F0, F0CV and IntCV for women (n = 16) and men (n = 16):  
 511 median values, z scores, significance tests (bold faces indicate statistically significant values) and effect size, *r*.  
 512

Acoustic characteristics	ADS	CDS	Wilcoxon signed rank test, <i>z</i>	<i>p</i>	Matched pair rank biserial correlation, <i>r</i>
<b>Women (n= 16)</b>					
Mean F0 (Hz)	<b>232.44</b>	<b>280.76</b>	<b>3.52</b>	<b>&lt; 0.001</b>	<b>1.000</b>
F0CV	0.088	0.051	-0.93	0.38	-0.265
IntCV	<b>0.87</b>	<b>0.98</b>	<b>2.02</b>	<b>0.044</b>	<b>0.574</b>
<b>Men (n=16)</b>					
Mean F0 (Hz)	<b>116.94</b>	<b>165.31</b>	<b>3.41</b>	<b>&lt; 0.001</b>	<b>0.971</b>
F0CV	0.037	0.056	1.19	0.25	0.338
IntCV	0.67	0.85	-0.34	0.74	-0.096

513

514

515

516 **Table 2** Results of LMM: effect of speech register (“ADS” or “CDS”), type of phrase (“do you want a treat”, “do you  
 517 want to play”, “see you later” and “how are you”) and interaction between the two factors on mean F0, F0CV and IntCV.  
 518 N = 15 women

519

Acoustic characteristics / Factors	Speech register	Type of phrase	Interaction
Mean F0	<b>&lt;0.001</b>	0.289	0.977
F0CV	0.472	<b>0.004</b>	0.708
IntCV	0.366	0.741	0.043

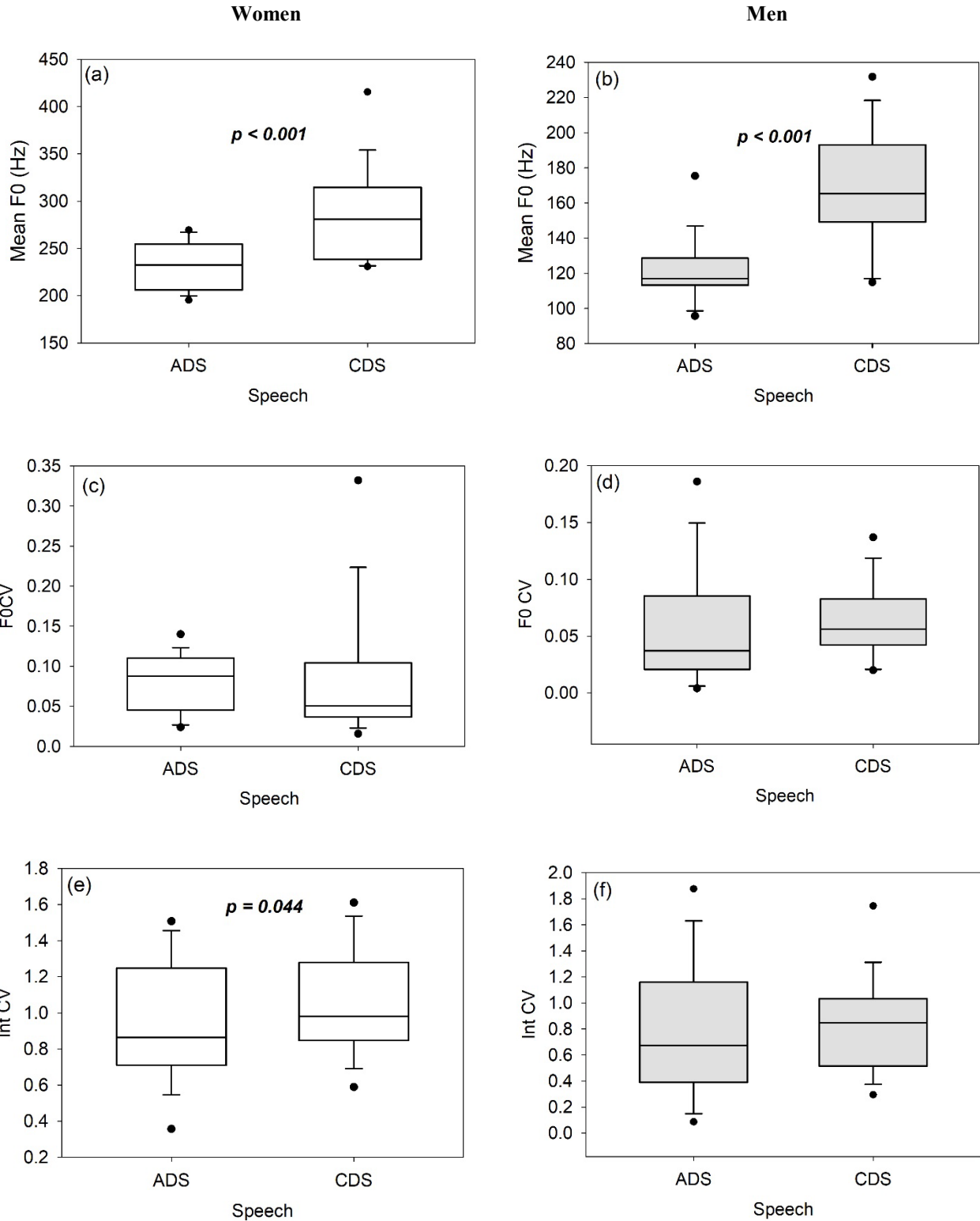
520

521

522

523 **Fig. 1** Comparison between “ADS” (telling the cat’s name to an adult) and “CDS” (calling a cat from a video) for mean  
 524 F0 (in Hz) of women (a) and men (b), for F0 CV of women (c) and men (d), for Int CV of women (e) and men (f). Median,  
 525 lower and upper quartile of the data are given, error bars represent the 10th and 90th percentiles; dots indicate 5th and  
 526 95th percentiles. Significant differences are indicated when present.  
 527

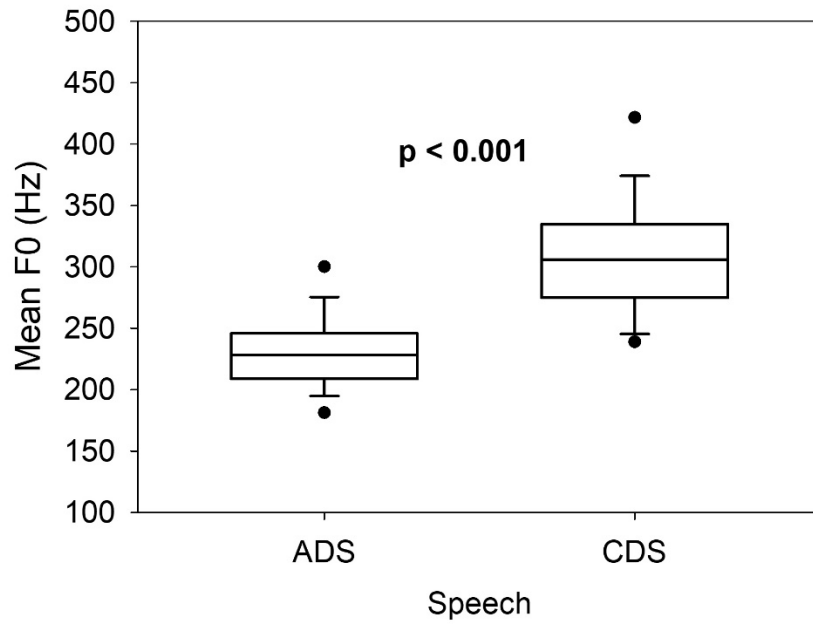
528  
 529



534  
 535  
 536



537 **Fig. 2** Effect of speech register on mean F0 (in Hz). Results for “ADS” and “CDS” of n = 15 women. Median, lower and  
538 upper quartile of the data are given, error bars represent the 10th and 90th percentiles; dots indicate 5th and 95th  
539 percentiles. Significant differences are indicated on the figure.



540

541

542