



The face behind the *caring* voice: A comparative study on facial prosodic features of dog-, infant- and adult-directed communication

Édua Koós-Hutás^{a,c,*}, Barbara A. Kovács^{b,c}, József Topál^{c,d}, Anna Gergely^{c,d}

^a Doctoral School of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

^b Doctoral School of Biology, ELTE Eötvös Loránd University, Budapest, Hungary

^c HUN-REN Research Centre for Natural Sciences, Institute of Cognitive Neuroscience and Psychology, Budapest, Hungary

^d NAP 3.0 Comparative Ethology Research Group

ARTICLE INFO

Keywords:

Doggerel

Motherese

Context-and species-specific facial prosody

Comparative cognition

ABSTRACT

Although it has been known for over 40 years that people tend to adopt similar communication styles with both infants and dogs, the visual component of infant-directed speech (motherese) has been overlooked in research. The aim of the current study was to investigate variations in visual prosodic features across different addressees (i.e. dogs, infants, and adults) and the intended goal of communication. Parents who were also dog owners interacted with their infant (6–18 m.o.), their family dog and their spouses as adult-directed control in three different, 1–2-minute-long scenarios: language tutoring, attention-getting, and fixed sentences. We comparatively analysed speakers' (N=46; 23 females, $M_{age}=34.41$ years, $SD=4.41$) facial expressions (i.e., happy, surprised, neutral) and emotional states (valence, arousal) across scenarios and conditions using Noldus FaceReader software. Results revealed a significant interaction effect of condition and situation. As expected, visual prosody was the least intense in the dog-directed condition, with language tutoring showing the lowest intensity. Fixed sentences, with lexical content intended for infants (i.e. nursery rhyme), exhibited consistently the highest intensity. Interestingly, adult-directed prosody toward spouses resembled infant-directed prosody more than expected. In summary, our results highlight disparities in the interpretation of facial expressions in both con- and interspecies communication. Despite the study's limitations, it emphasizes the importance of ongoing visual prosody characterisation, inviting further clarification and reconsideration of the dynamics between the visual and acoustic aspects of motherese and dog-directed talk as well as the role of bonding between interactants. These insights underscore the benefits of comparative cognition involving canines.

1. Introduction

Over 40 years ago, evidence supporting the existence of dog-directed (DD) speech or doggerel – a communication style resembling infant-directed (ID) talk – was first reported (Hirsh-Pasek and Treiman, 1982). Although different cultures and languages might result in variability in some properties (e.g., Farran et al., 2016) people can effectively recognise the unconsciously modulated ID speech register when hearing it and distinguish it from typical adult discourse (Grieser and Kuhl, 1988; Werker, Pegg and McLeod, 1994; Bryant and Barrett, 2007; Broesch and Bryant, 2015; Hilton et al., 2022). Recent demonstrations have revealed that employing a high-raised pitch with various pitch contours (resulting in a higher fundamental frequency – F0), a slower tempo and the use of simplified, frequently repeated words are tools

utilized by both mothers and fathers (e.g., Werker and McLeod., 1989; Gergely et al., 2017; 2023, Shapiro, Hippe and Ramírez, 2021). Compared to motherese, doggerel has an even higher F0 but lacks hyperarticulation (Gergely et al., 2017).

With respect to the main focus of our study, centred on the function of motherese and doggerel, a scientific consensus has not been reached. The vast majority of studies agree that language tutoring is one of the primary objectives of infant-directed speech (e.g. Burnham et al., 2002), but it also conveys information about the speaker's identity and intentions (Fernald, 1989; Mehler et al., 1978). Note that acquiring and maintaining the addressee's attention has also been suggested as another potential function (Fernald, 1985; Spinelli et al., 2017). Others argue that increasing positive interactions (Scherer, 1979; Murray and Trevarthen, 1985), expressing emotional state (Scherer, 1995) and

* Corresponding author at: Doctoral School of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary.

E-mail address: koos.edua@ppk.elte.hu (É. Koós-Hutás).

<https://doi.org/10.1016/j.applanim.2024.106203>

strengthening bonding between the interacting partners (Fernald, 1992) can also be linked to the affective function of motherese and doggerel.

An intriguing facet of recent literature is that while numerous publications delve into the acoustic features of motherese (for a review: Soderstrom, 2007; Cox et al., 2023) and doggerel (e.g. Ringrose, 2015) there is a notable scarcity of literature, if any, on their visual features. This imbalance may stem from the fact that, similar to dog-directed speech, ID talk heavily relies on the acoustic modality, promoting language acquisition, vowel discrimination (through hyperarticulation) and word segmentation in infants (e.g., Trainor and Desjardins, 2002; Song, Demuth and Morgan, 2010; Thiessen, Hill and Saffran, 2005; Wang et al., 2018). However, it is important to note that the perception of relevant visual cues during communication encompasses all kinds of non-verbal signals of body language: proximity, posture, gesture, mimicry (i.e., facial movement) and eye contact (e.g., Hadar, 1989) and the listener's linguistic ability level is one of the factors (e.g., gender, age, etc.) influencing the chosen speech style (e.g., DePaulo and Coleman, 1986; Kitamura and Burnham, 2003). Taken together, there is growing acceptance that both acoustic and visual modalities actively shape the dynamics of parent-infant interactions, contributing to the information content of infant-directed speech situations (Englund and Behne., 2020).

Despite its significance, investigating ID in a comprehensive, longitudinal manner is limited since infants and children display an exponential development curve compared to youth and adulthood (Craik and Bialystok., 2006). In line with the review by Weinstein and Baldwin (2022) which emphasised the need for further investigations into the multi-dimensional complexity of ID speech, incorporating experimental subjects with varying linguistic skills may help address some of the existing limitations in the literature.

Dogs are feasible candidates for such comparative studies (e.g. Gergely et al., 2023), not only because they are often regarded as “family members” (Smith, 1983; Tannen, 2004) and share the same social environment as infants, exhibiting child-like attachment behaviours (e.g., Topál et al., 1998), but also because they prefer speech with infant-directed talk properties (Burnham, 2002; Jeannin et al., 2017; Gergely et al., 2021) over casual adult-directed talk. This is unsurprising, given that they rely on human attentional cues such as the orientation of the body, head and eye (Gácsi et al., 2004) and they are able to detect humans' facial expressions (Müller et al., 2015) and emotions (Albuquerque et al., 2016). Moreover, eye contact with the human partner serves as a social reference for the dog and provides a rich source of information about human intentions and readiness to communicate (Savalli et al., 2016). At the same time, findings indicating the lack of vowel hyperarticulation in dog-directed speech (Gergely et al., 2017) suggest the absence of language tutoring function in such situations. This aligns with the concept of infants being treated as future speakers, in contrast to the non-speaker characteristic attributed to dogs.

Concerning the facial expressions employed by the human speaker in infant-directed speech, Chong et al. (2003) have described two primary facial expressions, characterised by exaggerations in the eye and mouth area, which express happy and surprised emotions. A more recent study (Gergely et al., 2023) aimed, to some extent, to replicate the findings of Chong et al. (2003) in a comparative framework and include dog-directed speech in the study. In a within-subject design, they manually coded still images extracted from videos using the gold standard method (i.e. Facial Action Coding System, FACS, Ekman and Friesen, 1978) and compared facial expressions and intensity levels across three situations: Task solving, Fixed sentences and Attention-getting (i.e., joint attention using an object). Authors identified six distinct facial expressions referred to as the ‘Prosodic Faces’. All six were commonly displayed when speaking to infants and dogs but were rarely observable or entirely absent when talking to friendly but unfamiliar adults. These results confirm the context-specificity of facial prosody. Another intriguing result of this study was the significantly less frequent and intense facial prosody with dogs compared to infants showed both by female and male speakers. This implies that facial

prosody might also have species-specific features.

The aim of the current study is to clarify the function of visual prosody in infant-directed speech through a comparative framework and methodological design that has proven successful, albeit with some necessary modifications. We examine facial prosodic features observed in adult-to-dog/infant/adult communicative interactions recorded in a natural, non-laboratory environment. Note, that in contrast to earlier studies, we employ a fine-grained resolution facilitated by the automated analysing program, the FaceReader software. In line with previously suggested limitations of using adult-directed speech toward unfamiliar adults as a control condition for infant-directed speech toward own infants (e.g. Trainor et al., 2000), we applied spouse-directed talk in the present study. This way, our three experimental conditions were more balanced for personal relationships between interactants, as speakers were talking to different members of the same family (i.e. family dog, infant and spouse). At the same time, it is important to note that measuring attachment styles and social status of the interactants was beyond the scope of this paper. Hence, conditions could vary in these parameters.

Regarding the question of whether specific facial features contribute to the affective or cognitive functions of visual prosody, we hypothesized that positive facial emotional expressions would be more intense and more frequent in ‘attention-getting’ compared to ‘language tutoring’ situations (for the description of different situations – see Methods). The intensity of visual prosody accompanying the ‘fixed sentences’ situation may fall between the intensities observed in the above two situations (Gergely et al., 2023). Furthermore, considering the significant differences in linguistic competence among dogs, infants, and adult partners, we anticipated the most pronounced addressee-specific distinctions in the language tutoring situation.

We also set out to investigate whether there are any addressee-specific facial prosodic features in adult-to-dog/infant/adult communication. In line with the findings of Gergely et al. (2023), we hypothesised a less emphasised visual prosody towards dogs compared to infants. Nevertheless, given that lover's talk shares acoustic properties with infant-directed speech (e.g., Kranjčić, 2016), we anticipated adult-directed communication to be closer to infant-directed communicative interactions in this regard.

2. Methods

2.1. Participants

Forty-six participants (23 women, 23 men, mean age $\pm$ SD = 34.41 $\pm$ 4.41 years) were recruited from 23 volunteering families. All participants were native Hungarian speakers. Having both a family dog (older than one year; 10 female and 22 male dogs, mean age $\pm$ SD: 5.98 $\pm$ 3.54 years) and an infant (6–18 months-olds, 11 girls and 12 boys, mean age $\pm$ SE: 10.42 $\pm$ 3.99 months) was a criterion for inclusion. Participants had no prior knowledge of the precise aim of the experiment, including details about facial expressions, hypotheses, and predictions. They were informed only that the study aimed to investigate infant-, dog-, and adult-directed communication.

2.2. Procedure

The data collection took place at the participants' homes in the presence of two experimenters. While one explained the instructions, the other experimenter managed the technical equipment required for recording. Each mother and father were separately recorded in a within-subject design while interacting with their infant (infant-directed condition, ID), their family dog (dog-directed condition, DD), and their spouses (adult-directed condition, AD). The participants were instructed to sit approximately 0.5–1 m from the addressee (Fig. 1). While leaning over or touching the addressee was not prohibited overall, participants were encouraged to keep a sitting position during the interaction and



Fig. 1. Examples of experimental arrangements. a) *infant-directed, ID*; b) *dog-directed, DD*; c) *adult-directed, AD* conditions. [a) and c) in *Language Tutoring*, b) in *Attention-getting situations*.] *Permission is obtained for the use of these participants' images for publication.

avoid covering their faces. The addressee remained in a relatively fixed, comfortable position, possibly at eye level with the speaker (or in a way that did not abrupt the recording of the speaker's face). Infants and dogs were gently held in the position by the other partner (helper) if needed (Fig. 1a). In these cases, the helper was passive and non-engaging. The interactions were recorded using three smartphones and one microphone. Each recording phase consisted of three situations (*attention-getting, language tutoring, and fixed sentences*) in the ID, DD, and AD conditions, respectively. The order of the situations was counter-balanced between speakers and conditions (i.e. AD, DD, ID), and it was semi-randomized (i.e. for each condition, one-third of speakers started with AT, one-third started with FIX and one-third with LT situations).

2.3. Attention-getting & maintaining (AT; 1 minute)

Participants were instructed to capture the addressee's attention and keep their focus for one minute. There were no restrictions on achieving this task; hence, the aim was to observe natural methods employed by parents to become the centre of the addressee's undivided attention. Therefore, no object was included, which was expected to result in more pronounced prosodic expressions.

2.4. Language tutoring (LG; 1+1 minutes)

This situation included habituation/training (object presenting) and testing (multiple-choice). Before the recording sessions, the experimenter selected one made-up word (label) out of three options ("danidu", "burida", "zibula") and two objects (a target and a non-target object) out of five (see Figure S1). The labels and objects were semi-randomised with the criteria that within each family, objects and labels selected for one parent could not be selected again for the other.

During the habituation phase, the participants helped the addressee build an association between the meaningless word (label) and the target object while simultaneously holding the two unknown objects in their hands. The participants could only refer to the distractor using demonstrative pronouns (i.e., 'this', 'that'). The instruction stated to speak about both objects separately for at least 30 seconds and use the target word/label as much as possible (at least thrice). The addressee was not allowed to touch either object during this phase.

In the second phase (after approximately 1 minute), the participants switched the objects in their hands and encouraged the addressee to find and choose the target object by referring to the previous label. If the addressee showed a choice by reaching or touching, the objects were switched back to the original position for repeated testing. The phase concluded after two clear choices, regardless of correctness. In the AD condition, participants hid their hands behind their backs while switching the objects' positions if desired. Partners were asked to guess the location of the target object (left or right hand).

2.5. Fixed sentences (FIX; max. 1 minute)

The participants were asked to tell a nursery rhyme: *Cini-cini muzsika; táncol a kis Zsuzsika; jobbra dől, balra dől; ticsök koma hegedül* (in English: "Cini-cini music plays; little Susan dances away; leaning to the right, leaning to the left; the cricket buddy plays the fiddle") and three predetermined sentences (from Gergely et al., 2017) to the addressee: (#1) *Nézd csak, milyen szép idő van odakint!* (In English: Just look! What nice weather!) (#2) *Akarsz sétálni egyet?* (In English: Do you want to go for a walk?) (#3) *Úgy látom, unatkozol. Nem csinálunk valami mást?* (In English: You seem really bored. Shouldn't we do something else?).

2.6. Stimuli and materials

3 smartphones (iPhone 8 S/ Samsung Galaxy A22/ iPhone 13 pro, Samsung Galaxy A70, OnePlus Nord N10) with additional tripod stands and a microphone (Zoom F2 recorder with LMF-2 lavalier microport) were used for recording. The first camera - always the OnePlus Nord N10 [resolution: 1080×60 fps; refresh rate: 90 Hz] - was focused on the speaker's torso. Since this smartphone was held by one experimenter, flexibly following the natural movement of the participant was applicable. The second camera recorded the reaction of the addressee, while the whole interaction (from the profile side of the communicating partners) was captured by the third smartphone with a wide angle. The microphone was clipped on the speaker's collar or sleeve.

The nursery rhyme and the three fixed sentences used in the attention-getting situation were printed on an A/4 sheet of paper, with a font size visible from a 2–3-meter distance, to prevent confusion and disturbance. The set of eligible objects used in the language tutoring situation consisted of a green silicone spoon rest, a yellow silicone tea infuser, 1/5th of a pink silicone waffle mould, a lavender plastic garlic press and an orange silicone door/window stopper (see Figure S1).

2.7. Data acquisition and analyses

First, video recordings of the speakers' faces were analysed using the FaceReader software (Noldus FaceReader 8, www.noldus.com/facereader; Skiendziel et al., 2019; Kyranides et al., 2022) based on the Facial Action Coding System method (FACS; Ekman and Friesen, 1978), facilitating a quicker, more reliable analysis and allowing information-gathering from continuous variables. One video consisted of all three situations, resulting in 3 videos per speaker. The starting and ending points of situations as well as the non-relevant parts of the recordings (transition between situations or where the situation got disrupted due to, e.g., the speaker forgetting instructions and addressing the experimenter), were manually coded. The clips falling into the latter category were excluded, while the software automatically coded certain sections unsuitable for evaluation (e.g., when the face was partially or

covered entirely) (i.e., FIT_FAILED). These missing values (NA) stayed a part of the data set. The analysis focused on the participants' (speakers') facial prosodic features and incorporated five estimated intensity levels (0–1) per second for seven emotional faces (happy, neutral, surprised, angry, sad, scared, disgusted), two emotional states (valence and arousal) and 20 Action Units (i.e., individual components of muscle movement of the face) To remain in the scope of the research questions, the following statistical examinations focused on the two most relevant emotional expressions previously reported (Chong et al., 2003; Gergely et al., 2023) to dominate these interactions (Happy, Surprised) and two emotional states (Valence, Arousal). In addition, we also analysed Neutral faces as a contrast to the emotionally rich faces. Arousal indicates the subject's activity level, that is, whether any facial movement enables the appearance of facial expressions. Valence implies the subject's positive or negative emotional state and is linked to the fundamental frequency (F0), an essential primer acoustic characteristic of ID speech (Dmitrieva et al., 2008). It is continuously calculated by subtracting the highest intensity among the three negative expressions (i.e., "angry", "sad", and "disgusted") from the intensity of 'Happy' at a given moment (Loijens and Krips, 2018, p. 7).

R studio (<https://www.rstudio.com/>) was used for the statistical analysis (R 4.2.3 using RStudio 2023.12.0+396). We analysed the intensity of facial prosodic features by conditions and situations with linear mixed models (GLMM and LMM, nlme and lme4 package, glmer and lme functions, Bates et al., 2015; Pinheiro and Bates, 2000) with AIC-based backwards elimination (MASS package, drop1 function, Venables and Ripley, 2002) to find parsimonious models. We analysed the distribution of each variable with histograms and QQ plots. In line with the distributions, we applied GLMM for Gamma distribution with log link function for 'Happy' and 'Surprised' facial expressions, GLMM for Gaussian distribution with log link function for 'Neutral' and 'Arousal', and LMM fit by maximum likelihood for 'Valence'. As recent findings suggest no major prosodic differences based on the speaker's gender (Gergely et al., 2017; Gergely et al., 2023) and our hypotheses were not specified for gender effects, we did not include speaker's gender as a fixed effect to the models. Therefore, all models included condition (AD, DD, ID), situation (AT, LG, FIX) and their interaction as fixed effects. First, we included speaker ID and family ID in the models as

a random intercept (speaker nested in family) in order to control for dependence and repeated measurements. After comparing model performance (compare performance function) and checking the explained variance, family ID was dropped out as it explained no variance and only speaker ID was included as a random intercept in all final models. For post-hoc pairwise comparisons, we used Tukey method (emmeans package, Lenth, 2023).

3. Results

3.1. Emotional facial expressions

Model selection showed a significant interaction effect of condition $\times$ situation for *Happy* facial expression (LRT: $\chi^2_4=633.54$, $p<0.001$). Pairwise comparisons revealed that *Happy* expressions appeared in AD and ID speech with significantly higher intensity than in DD speech in all situations (all $p<0.001$, Fig. 2, for detailed statistics, see Table S1). While *Happy* facial expression intensity differences varied across situations between ID and AD conditions. In FIX situation, participants displayed similarly intense *Happy* faces toward their infants and spouses ($p=0.494$, Fig. 2). During AT situation, they displayed more intense *Happy* expressions in ID than in AD condition ($p<0.001$, Fig. 2). On the contrary, during LG situation, more intense *Happy* expressions were found in AD than in ID condition ($p<0.001$, Fig. 2). Within conditions, differences in the *Happy* expression intensities between situations followed a similar pattern in ID and AD conditions, namely that speakers used the highest intensity during FIX, then in AT situation, and the lowest intensity was found during LG situation (all $p<0.001$, Fig. 2, for detailed statistics, see Table S1). In DD condition, however, the lowest intensity for *Happy* expressions was found during AT, then in LG and the highest in FIX situation (all $p<0.001$, Fig. 2, for detailed statistics, see Table S1).

In sum, *Happy* facial expressions were the most intense toward conspecific partners (i.e. infants and spouses) and FIX sentences situation evoked the most intense *Happy* faces, which included the nursery rhyme (see Table 1).

Similar to *Happy*, the GLMM showed significant condition $\times$ situation interaction for *Surprised* facial expression (LRT: $\chi^2_4=226.17$,

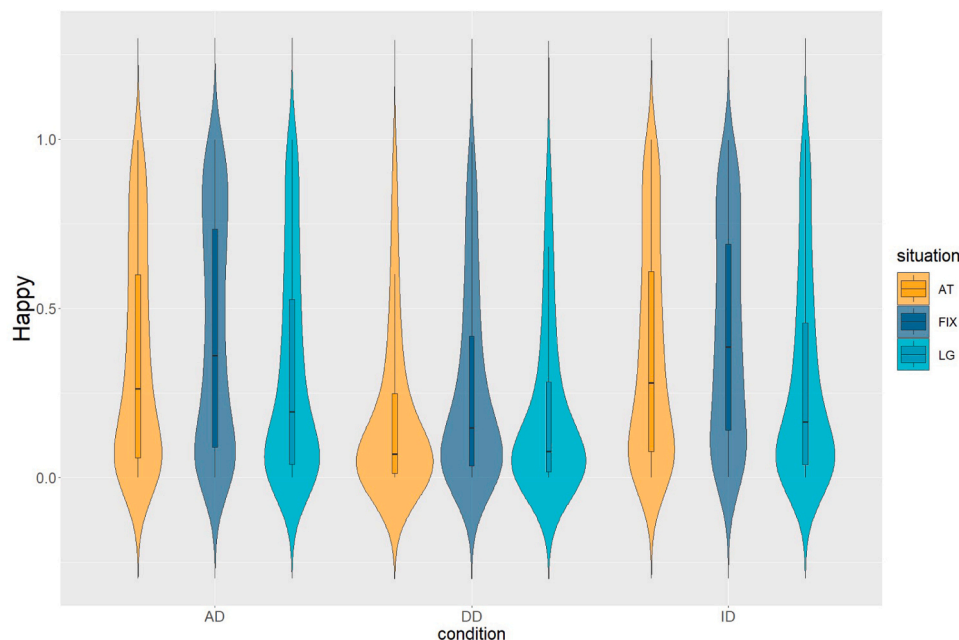


Fig. 2. The intensity level of *Happy* facial prosodic features during situations across conditions. [Within the boxplots, the horizontal line represents the median, the box shows the quartiles, the whiskers represent the range, and the violin plots depict the density of the data. Situations yellow - attention-getting; dark blue - fixed sentences and nursery rhymes; turquoise - language tutoring. Conditions: AD - adult-directed; DD - dog-directed; ID - infant-directed.].

Table 1

Summary of results on facial prosody toward adults (AD, i.e. spouses), infants (ID) and dogs (DD) in three situations, including Attention-getting (AT), Language tutoring (LG) and Fixed sentences (FIX) with nursery rhymes.

	Happy	Surprised	Neutral	Valence	Arousal
Attention-getting	(AD<ID)>DD	(AD>ID)>DD	(DD>AD)>ID	(AD>ID)>DD	(AD<ID)>DD
Language Tutoring	(AD>ID)>DD				(AD>ID)>DD
Fix sentences	(AD=ID)>DD				(AD=ID)>DD
Adult-directed	(FIX>AT)>LG	(FIX>AT)>LG	(AT=LG)>FIX	(FIX>AT)>LG	(FIX>AT)>LG
Infant-directed			LG>(AT=FIX)		
Dog-directed	(FIX>LG)>AT		(AT>LG)>FIX	FIX>(AT=LG)	

$p<0.001$). Comparing conditions pairwise, *Surprised* expressions were the most intense during AD, then in ID, and the least intense during DD condition in all three situations (all $p<0.001$, Fig. 3, for detailed statistics, see Table S2). FIX again was the leading situation with the highest intensity of *Surprised* expressions in AT, and even less intensity was observed during LG situation toward all partners (all $p<0.001$, Fig. 3, for detailed statistics, see Table S2).

Overall, *Surprised* faces of the highest intensity were found towards adults (i.e. the spouses), especially when participants were reciting fixed sentences including a nursery rhyme (see Table 1).

Model selection again showed a significant effect of condition $\times$ situation for *Neutral* facial expressions ($LRT: \chi^2_4=994.48, p<0.001$), though it displayed an opposite tendency to *Happy* and *Surprised*. The post hoc pairwise comparison shed light on significant differences between conditions: the least amount of facial expressions lacking emotional intensity (i.e.: *Neutral* expressions of lowest intensity) was found in ID, while the most intense *Neutral* faces appeared during DD speech, and AD fell between them (all $p<0.001$, Fig. 4, for detailed statistics, see Table S3). During DD condition, the most neutral facial expressions were present in AT, followed by LG, and last came FIX with the least intense *Neutral* faces (all $p<0.001$). Although a similar tendency was shown in AD, namely FIX possessing the least intense *Neutral* expressions, AT and LG had no significant difference ($p=0.574$, Fig. 4). In the case of ID speech, participants displayed the most *Neutral* faces in LG, while AT and FIX did not differ significantly ($p=0.964$, Fig. 4).

To conclude, the most *Neutral* facial expressions appeared towards

dogs, and the effect of the situation varied according to the partner (see Table 1).

3.2. Emotional state during situations

Concerning the *Valence*, LMM revealed a significant interaction effect of condition $\times$ situation ($LRT: \chi^2_4=587.89, p<0.001$), with the post hoc comparison showing significant differences between all conditions (all $p<0.001$, Fig. 5, for detailed statistics, see Table S4). During all three situations, speakers expressed the most positive facial expressions toward adults (i.e. their spouses) and then toward their infants and the least positive faces were expressed toward dogs (all $p<0.001$, Fig. 5, for detailed statistics, see Table S4). Overall, FIX was the leading situation in all conditions with the most positive valence, followed by AT and LG presented the least positive valence (all $p<0.001$, except AT vs. LG in DD condition, $p=0.217$). It is important to note that speakers' facial emotional states toward dogs were close to zero (i.e., either no emotions were shown, or if any positive or negative emotions appeared on the face, their intensity balanced out). In sum, facial emotional state (i.e. *Valence*) was the most positive towards adults (i.e. spouses) and FIX situation, again, evoked the most emotional and positive faces (see Table 1).

Finally, the GLMM model indicated a significant interaction effect of condition $\times$ situation for *Arousal* ($LRT: \chi^2_4=490.17, p<0.001$). The lowest level of facial muscle movements (i.e. *Arousal*) was observed in DD as opposed to ID and AD conditions in all three situations (all $p<0.001$, Fig. 6, for detailed statistics, see Table S5). Differences in

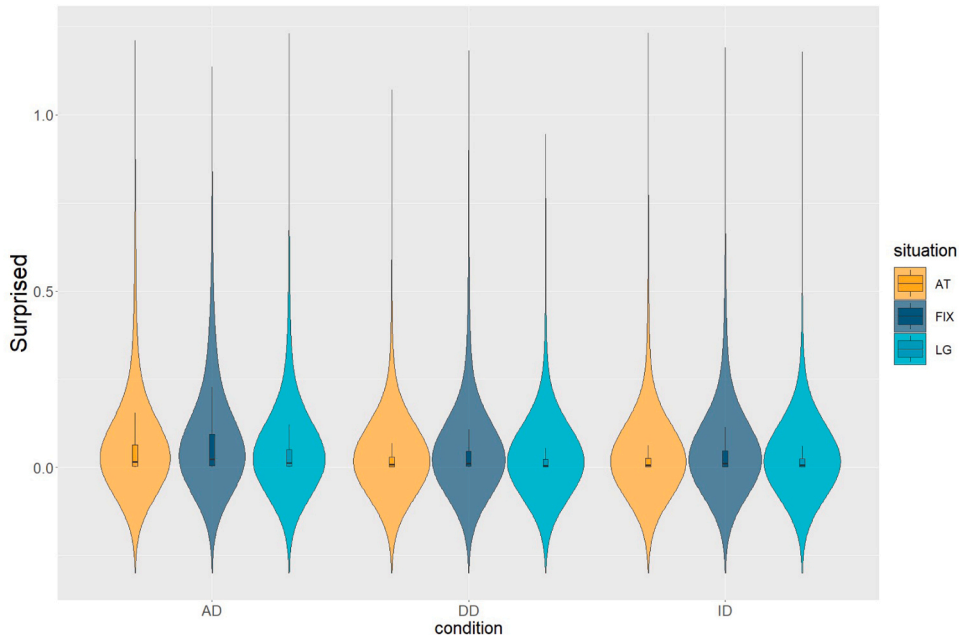


Fig. 3. The intensity level of *Surprised* facial expression during situations across conditions. [Within the boxplots, the horizontal line represents the median, the box shows the quartiles, the whiskers represent the range, and the violin plots depict the density of the data. Situations: yellow - attention-getting; dark blue - fixed sentences and nursery rhymes; turquoise - language tutoring. Conditions: AD - adult-directed; DD - dog-directed; ID - infant-directed.].

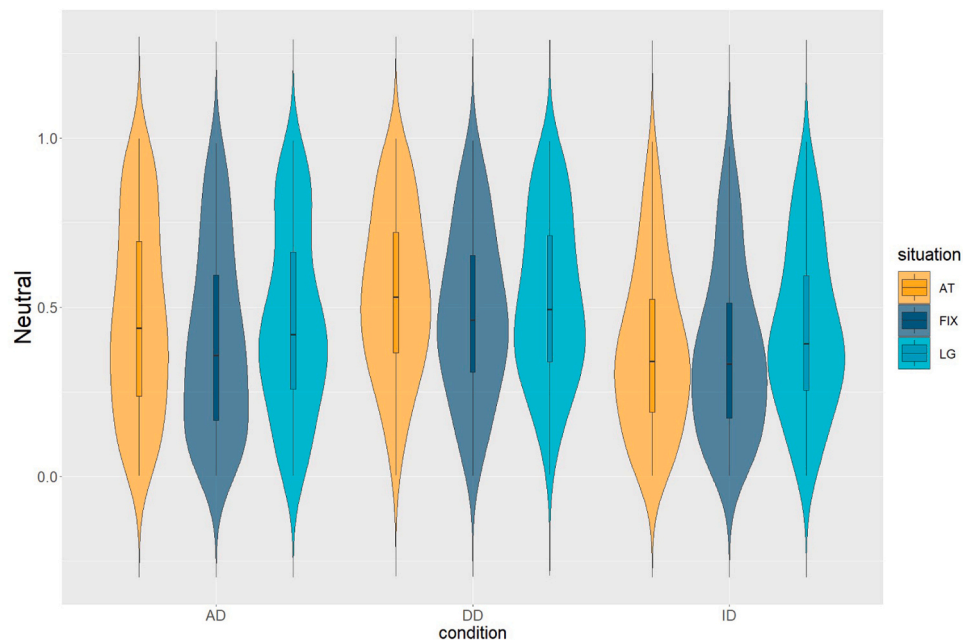


Fig. 4. The intensity level of *Neutral* facial expression during situations across conditions. [Within the boxplots, the horizontal line represents the median, the box shows the quartiles, the whiskers represent the range, and the violin plots depict the density of the data. Situations: yellow - attention-getting; dark blue - fixed sentences and nursery rhymes; turquoise - language tutoring. Conditions: AD - adult-directed; DD - dog-directed; ID - infant-directed.].

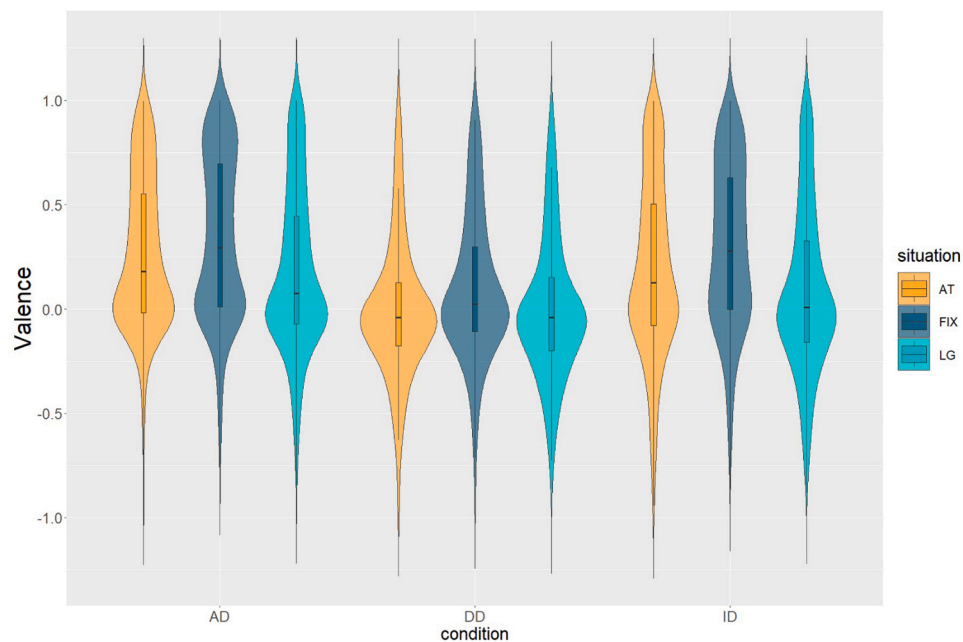


Fig. 5. Emotional state measured on the face by *Valence* during situations across conditions. [Within the boxplots, the horizontal line represents the median, the box shows the quartiles, the whiskers represent the range, and the violin plots depict the density of the data. Situations: yellow - attention-getting; dark blue - fixed sentences and nursery rhymes; turquoise - language tutoring. Conditions: AD - adult-directed; DD - dog-directed; ID - infant-directed.].

facial *Arousal* toward infants and adults varied across situations. During AT situation, a higher level of *Arousal* was present towards infants, then towards adults ($p < 0.001$, Fig. 6, for detailed statistics, see Table S5). In contrast, during LG situation, participants displayed significantly more muscle movements on their faces towards adults than toward infants ($p < 0.001$), while during FIX situation no difference was found between ID and AD conditions ($p = 0.819$, Fig. 6, for detailed statistics, see Table S5). Pairwise comparisons revealed that participants exhibited significantly more facial muscle movements during FIX situations than in AT, and significantly fewer facial movements were observed during

LG situations toward all partners (all $p < 0.001$, Fig. 6, for detailed statistics, see Table S5; Table 1).

4. Discussion

One of the initial objectives of the study was to improve the methodology used by Gergely et al. (2023) and to compare the nature of facial prosodic features in dog-, infant- and adult-directed talk within cognitive (Language tutoring) and attentive/affective (Attention-getting and Fixed sentences) scenarios. At first glance, the results confirm the

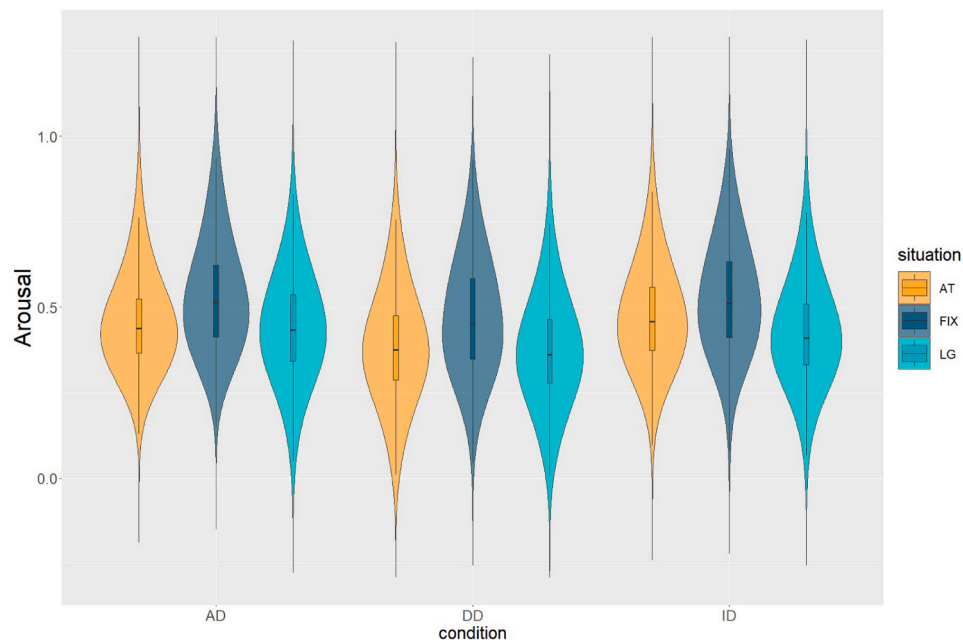


Fig. 6. The intensity level of *Arousal* of the facial prosodic features during situations across conditions. [Within the boxplots, the horizontal line represents the median, the box shows the quartiles, the whiskers represent the range, and the violin plots depict the density of the data. Situations: yellow - *attention-getting*; dark blue - *fixed sentences and nursery rhymes*; turquoise - *language tutoring*. Conditions: AD - *adult-directed*; DD - *dog-directed*; ID - *infant-directed*.].

hypothesised expectation that communicating positive content intended for infants is associated with more pronounced facial muscle movement. At the same time, a somewhat surprising twist – namely, the low intensity of facial prosody toward dogs in contrast to the almost infant-directed-like intensity toward adults (i.e. spouses) and the exaggerated facial prosody during *Fixed sentences* situation – refreshes the interpretation of expectations.

Let us systematically address our questions and inspect the findings point by point: As for function dependency, our results highlight that when conveying a positive emotional state is highly relevant (i.e., in the case of affective functions), the speaker's face serves as an excellent medium for reflecting e.g., their engagement and inner joy. The increased facial intensity observed in the *Fixed sentences* (nursery rhymes) and *Attention-getting* situations as well as the decreased intensity during *Language tutoring* toward conspecific partners (i.e. adult-directed and infant-directed conditions) aligns with our hypothesis. This indicates that exaggerated visual (facial) expressions serve as an efficient tool for capturing and maintaining attention during conspecific communicative interactions. These facial expressions may function as mediators in joint attentional situations, serving both as initiator and feedback to signal the speaker's readiness and willingness to engage in interactions with either an adult or a child partner. Moreover, our results revealed a high impact of infant-directedness in a given situation on facial prosody. Namely, a high level of intense facial prosody during the *Fixed sentences* situation (involving a popular nursery rhyme) was clearly perceived by the speakers as content intended for infants or children. The speech style of such rhymes is usually performed with distinctive rhythm and melody, which facilitate language acquisition (e.g. Suppa-nen et al., 2019). When uttering rhymes, speakers in the present study seem to switch to this unique speech style, regardless of whether the addressee is an infant, an adult, or a dog. While it has been shown that acoustic prosody is strongly associated with the infant-directedness of a situation (e.g. Falk and Audibert, 2021; Hilton et al., 2022), our results provide evidence that facial prosody operates similarly, and a high level of infant-directedness of the given situation has the potential to elicit exaggerated facial movements. However, it is worth mentioning that the nursery rhyme was followed by three predetermined sentences in the *Fixed sentences* situation. Thus, the high level of intense facial prosody

observed even during the recitation of fixed sentences might result from the priming effect of the nursery rhyme.

A recent comparative study on visual prosody revealed similar facial expressions toward infants and dogs but these were displayed less frequently with lower intensity toward dogs compared to infants (Gergely et al., 2023). It has been suggested that the similarity between dogs and infants is linked to their comparable emotional needs, which are, for example, evident in their attentiveness to positive facial expressions. However, differences in frequency and intensity may be attributed to their distinct cognitive capabilities (e.g., differences in understanding lexical content, see Gergely et al., 2023). Based on this, we expected that ID and DD would be similar in terms of facial movements during affective situations (i.e. Attention-getting and Fixed sentences scenarios) while differences were expected during more cognitive situations (i.e. Language tutoring). On the contrary, we found significant differences with generally more intense facial prosody toward infants than to dogs in all situations. Even in the *Fixed sentences* situation, characterized by intense facial prosody and a high level of infant directedness, fewer facial movements were evoked when speakers addressed their dogs. Furthermore, the analyses of *Arousal* and *Valence* clearly indicate that speakers' facial movements were more active and reflected a more positive emotional state when communicating with another human family member (i.e., their spouses and their infants) than with their own family dogs.

An examination of the values of the *Neutral* state (c.f. Fig. 4.) confirms that the owner's face generally remains rather uninvolved during dog-directed talk. It is important to discuss differences between human and canine facial communicative cues in order to conclude this phenomenon. While showing teeth and gum is associated with social smiles and happy emotions in humans, it is linked to agonistic and aggressive encounters in dogs and wolves (e.g. Ekman, 1982; Mech and Boitani, 2010; Siniscalchi et al., 2018). Wide-opened eyes with visible sclera also express different emotional states in the two species: surprise and attention in humans, whereas fear and anxiety in dogs (e.g. Ekman, 1982; Siniscalchi et al., 2018). Based on these, we can assume that dogs might respond with avoidance and would show stress signals to intense, infant-directed-like facial prosody in such close face-to-face scenarios. In line with this, it is probable that speakers reduce the intensity of facial

prosody during interspecies communication with dogs. Exaggerated acoustic prosody toward dogs, which can be even more intense than baby talk (e.g., Gergely et al., 2017), might serve as a primary prosodic modality in dog-directed communication, compensating for the reduced facial prosody. Future studies are needed to confirm these assumptions, focusing on investigating dogs' reactions and behaviour in response to intense facial versus acoustic prosody exhibited by humans.

Going further into the addressee-specific nature of facial prosodic features, we should discuss the unexpected pattern of adult-directed communication, specifically measured toward own spouses in the present study. Since speaking with a stranger is never equivalent to addressing one's own child and dog in terms of familiarity, emotional closeness, and intimacy (e.g., Morton, 1978), it has been suggested that AD would serve as a more viable control for ID speech if bonding between the interactants were controlled (Trainor et al., 2000). When it comes to facial prosody, we can conclude that personal relationships with the adult partner had a significant impact on the results. Namely, in previous studies, AD condition (i.e. toward nice but unfamiliar adults) remained below ID and DD condition in terms of the frequency and intensity of facial movements (e.g. Chong et al., 2003; Gergely et al., 2023).

While one of the most recognized functions of prosody is to enhance bonding between participants, there are only a few studies on how the presence or absence of bonding and/or attachment affects acoustic and visual prosody. In conclusion, the findings of our study provide evidence that facial prosody is affected by context and has species-specific aspects, as well as highlight the importance of conducting further investigations that include controls for personal relationships between the interlocutors across conditions.

Nonetheless, our findings have certain limitations. Although the goal of the rather bold decision to place the observations at the participants' homes was to capture the true nature of the infant-directed talk, stepping outside of an easily controllable environment introduces the possibility of artefacts. According to Cooper and Aslin (1989), while naturalistic observation in home settings takes social context into account and provides a more comprehensive understanding of the complexity of infant-directed talk (Ramírez-Esparza et al., 2014), the potentially distorting effects of the artificial environment cannot be entirely ruled out, influenced by recording equipment (e.g., Ozturk and Hill, 2018) and the presence of experimenters (Tamis-LeMonda et al., 2017). However, it is worth noting that by not requiring a suspension of everyday tasks and daily routines to achieve stillness and quiet in the household, all the data obtained reflect the naturally noisy environment to some extent. Additionally, a noteworthy limitation is that our study was conducted in a Hungarian language, middle-class environment. Social norms of appropriate behaviour in different situations might dictate various patterns of infant-, dog- and even adult-directed speech across cultures. For further research, it would be beneficial to incorporate the speakers' individual differences (e.g. personality traits). Furthermore, the quality of attachment should also be considered, as Marinelli et al. (2007) suggest the potential for competing nature of attachment when a dog owner becomes a parent. Whilst the included situations held different ecological relevance, we believe they are viable replicas of everyday challenges faced by both parents and dog owners throughout life. Their varying relevance proves to be an asset rather than a hindrance in elucidating the underlying functions, mechanism, and initiatives of these specific forms of communication. Nevertheless, understanding how speakers choose to overcome these challenges or interpret the goals of interaction provide valuable insights into both inter- and conspecific communication.

Funding

This research was funded by the Hungarian Scientific Research Fund (NKFIH grant no. FK142968) and by the Hungarian Brain Research Program (HBRP) 3.0 NAP.

Ethical statement

The study was conducted with the approval of the United Ethical Review Committee for Research in Psychology (EPKEB; No. 2022–85). The adult human participants provided their written informed consent to participate in this study. Written informed consent was obtained from the adult participants (parents of infants) to publish any identifiable images or data in this article.

CRediT authorship contribution statement

Édua Koós-Hutás: Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Barbara A. Kovács:** Investigation, Data curation. **József Topál:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Anna Gergely:** Writing – review & editing, Visualization, Supervision, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We are grateful to the participating families and to Anna Dallos and Mandula Koós-Hutás for their help in data acquisition.

Supplementary Material and Data Availability

Detailed statistics are available in the [Supplement Material](#) (Koos-Hutás et al. Supplement). Raw data and R script are available at Mendeley Data (<https://data.mendeley.com/preview/rzmxrjhtdf> <https://data.mendeley.com/datasets/rzmxrjhtdf/2>).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.applanim.2024.106203](https://doi.org/10.1016/j.applanim.2024.106203).

References

- Albuquerque, N., Guo, K., Wilkinson, A., Savalli, C., Otta, E., Mills, D., 2016. Dogs recognize dog and human emotions. *Biol. Lett.* 12 (1), 20150883.
- Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed effects models using LME4. *J. Stat. Softw.* 67, 1–48.
- Broesch, T.L., Bryant, G.A., 2015. Prosody in infant-directed speech is similar across western and traditional cultures. *J. Cogn. Dev.* 16 (1), 31–43.
- Bryant, G.A., Barrett, H.C., 2007. Recognizing Intentions in Infant-Directed Speech: Evidence for Universals. *Psychol. Sci.* 18 (8), 746–751. <https://doi.org/10.1111/j.1467-9280.2007.01970.x>.
- Burnham, D., Kitamura, C., Vollmer-Conna, U., 2002. What's new, pussycat? On talking to babies and animals. *Science* 296 (5572), 1435.
- Chong, S.C.F., Werker, J.F., Russell, J.A.C., 2003. Three Facial Expressions Mothers Direct to Their Infants. *Infant Child Dev.* 12 (3), 211–232.
- Cooper, R.P., Aslin, R.N., 1989. The language environment of the young infant: implications for early perceptual development. *Can. J. Psychol. / Rev. Can. De Psychol.* 43 (2), 247.
- Cox, C., Bergmann, C., Fowler, E., Keren-Portnoy, T., Roepstorff, A., Bryant, G., Fusaroli, R., 2023. A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nat. Hum. Behav.* 7 (1), 114–133. <https://doi.org/10.1038/s41562-022-01452-1>.
- Craik, F.I.M., Bialystok, E., 2006. Cognition through the lifespan: mechanisms of change. *Trends Cogn. Sci.* 10 (3), 131–138. <https://doi.org/10.1016/j.tics.2006.01.007>.
- DePaulo, B.M., Coleman, L.M., 1986. Talking to children, foreigners, and retarded adults. *J. Personal. Soc. Psychol.* 51 (5), 945–959. <https://doi.org/10.1037/0022-3514.51.5.945>.
- Dmitrieva, E.S., Gel'man, V.Y., Zaitseva, K.A., Orlov, A.M., 2008. Dependence of the perception of emotional information of speech on the acoustic parameters of the stimulus in children of various ages. *Hum. Physiol.* 34, 527–531.

- Ekman, P., 1982. *Emotion in the Human Face*. Cambridge University Press.
- Ekman, P. & Friesen, W.V. (1978). *Facial Action Coding System: Investigator's Guide Part Two: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press: Palo Alto, CA.
- Englund, N., Behne, D.M., 2020. Perception of audiovisual infant directed speech. *Scand. J. Psychol.* 61 (2), 218–226.
- Falk, S., Audibert, N., 2021. Acoustic signatures of communicative dimensions in codified mother-infant interactions. *J. Acoust. Soc. Am.* 150 (6), 4429–4437. <https://doi.org/10.1121/10.0008977>.
- Farran, L.K., Lee, C.C., Yoo, H., Oller, D.K., 2016. Cross-cultural register differences in infant-directed speech: An initial study. *PloS One* 11 (3), e0151518. <https://doi.org/10.1371/journal.pone.0151518>.
- Fernald, A., 1985. Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* 8 (2), 181–195.
- Fernald, A., 1989. Intonation and communicative intent in mothers' speech to infants: is the melody the message? *Child Dev.* 60 (6), 1497–1510.
- Fernald, A., 1992. Meaningful melodies in mother's speech to infants. In: Hanus-Papousek, U.J., Papousek, M. (Eds.), *Nonverbal Vocal Communications*. Cambridge University Press, New York.
- Gácsi, M., Miklósi, Á., Varga, O., Topál, J., Csányi, V., 2004. Are readers of our face readers of our minds? Dogs (*Canis familiaris*) show situation-dependent recognition of human's attention. *Anim. Cogn.* 7, 144–153. <https://doi.org/10.1007/s10071-003-0205-8>.
- Gergely, A., Faragó, T., Galambos, Á., Topál, J., 2017. Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: an acoustic analysis. *Sci. Rep.* 7, 13739 <https://doi.org/10.1038/s41598-017-13883-2>.
- Gergely, A., Tóth, K., Faragó, T., Topál, J., 2021. Is it all about the pitch? Acoustic determinants of dog-directed speech preference in domestic dogs, *Canis familiaris*. *Anim. Behav.* 176, 167–174. <https://doi.org/10.1016/j.anbehav.2021.04.008>.
- Gergely, A., Koós-Hutás, É., Filep, L.A., Kis, A., Topál, J., 2023. Six facial prosodic expressions caregivers similarly display to infants and dogs. *Sci. Rep.* 13 (1), 929. <https://doi.org/10.1038/s41598-022-26981-7>.
- Grieser, D.L., Kuhl, P.K., 1988. Maternal speech to infants in a tonal language: Support for universal prosodic features in motherese. *Dev. Psychol.* 24 (1), 14.
- Hadar, U., 1989. Gestural modulation of speech production: the role of head movement. *Lang. Commun.* 9 (4), 245–257.
- Hilton, C.B., Moser, C.J., Bertolo, M., Lee-Rubin, H., Amir, D., Bainbridge, C.M., Simson, J., Knox, D., Glowacki, L., Alemu, E., Galbarczyk, A., Jasienska, G., Ross, C. T., Neff, M.B., Martin, A., Cirelli, L.K., Trehub, S.E., Song, J., Kim, M., Schachner, A., Vardy, T.A., Atkinson, Q.D., Salenius, A., Andelin, J., Antfolk, J., Madhivanan, P., Siddaiah, A., Placek, C.D., Salali, G.D., Keestra, S., Singh, M., Collins, S.A., Patton, J. Q., Scaff, C., Stieglitz, J., Cutipa, S.C., Moya, C., Sagar, R.R., Aanyawire, M., Mabulla, A., Wood, B.M., Krasnow, M.M., Mehr, S.A., 2022. Acoustic regularities in infant-directed speech and song across cultures. *Nat. Hum. Behav.* 6 (11), 1545–1556. <https://doi.org/10.1038/s41562-022-01410-x>.
- Hirsh-Pasek, K., Treiman, R., 1982. Doggerel: motherese in a new context. *J. Child Lang.* 9, 31. <https://doi.org/10.1017/S0305000900003731>.
- Jeannin, S., Gilbert, C., Amy, M., Leboucher, G., 2017. Pet-directed speech draws adult dogs' attention more efficiently than Adult-directed speech. *Scientific reports* 7 (1), 4980.
- Kitamura, C., Burnham, D., 2003. Pitch and communicative intent in mother's speech: adjustments for age and sex in the first year. *Infancy* 4 (1), 85–110. https://doi.org/10.1207/S15327078IN0401_5.
- Kranjčić, T. (2016). *Baby talk in adult-to-adult communication* (Doctoral dissertation).
- Kyranides, M.N., Petridou, M., Gokani, H.A., Hill, S., Fanti, K.A., 2022. Reading and reacting to faces, the effect of facial mimicry in improving facial emotion recognition in individuals with antisocial behavior and psychopathic traits. *Curr. Psychol.* 1–14. <https://doi.org/10.1007/s12144-022-02749-0>.
- Lenth, R.V. (2023) emmeans: estimated marginal means, aka least-squares means. R package version 1.8.5. (<https://CRAN.R-project.org/package=emmeans>).
- Loijens, L., & Krips, O. (2018). FaceReader methodology note. A white paper by Noldus Information Technology. Retrieved from (https://info.noldus.com/hubfs/resources/noldus-white-paper-facereader-methodology.pdf?utm_campaign=
- Marinelli, L., Adamelli, S., Normando, S., Bono, G., 2007. Quality of life of the pet dog: Influence of owner and dog's characteristics. *Appl. Anim. Behav. Sci.* 108 (1–2), 143–156. <https://doi.org/10.1016/j.applanim.2006.11.018>.
- Mech, L.D., Boitani, L., 2010. *Wolves: Ecology, Behavior and Conservation*. University of Chicago Press, Chicago.
- Mehler, J., Bertoncini, J., Barriere, M., Jassik-Gerschenfeld, D., 1978. Infant recognition of mother's voice. *Perception* 7, 491–497.
- Morton, T.L., 1978. Intimacy and reciprocity of exchange: a comparison of spouses and strangers. *J. Personal. Soc. Psychol.* 36 (1), 72–81. <https://doi.org/10.1037/0022-3514.36.1.72>.
- Müller, C.A., Schmitt, K., Barber, A.L.A., Huber, L., 2015. Dogs can discriminate emotional expressions of human faces. *Curr. Biol.* 25 (5), 1. <https://doi.org/10.1016/j.cub.2014.12.055>.
- Murray, L., Trevarthen, C., 1985. Emotional regulation of interactions between two-month-olds and their mothers. In: Field, T.M., Fox, N.A. (Eds.), *Social perception in infants*. Ablex Publishers, Norwood, NJ, pp. 177–197.
- Ozturk, G., & Hill, S. (2018). Mother-child interactions during shared reading with digital and print books. *Early Child Development and Care*.
- Pinheiro, J.C., Bates, D.M., 2000. *Mixed-Effects Models in S and S-PLUS*. Springer. <https://doi.org/10.1007/b98882>.
- Ramírez-Esparza, N., García-Sierra, A., Kuhl, P.K., 2014. Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev. Sci.* 17 (6), 880–891.
- Ringrose, C., 2015. Pitch change in dog-directed speech. *Lifespan. Styles* 1 (4).
- Savalli, C., Resende, B., Gaunet, F., 2016. Eye contact is crucial for referential communication in pet dogs. *PLoS ONE* 11 (9), e0162161. <https://doi.org/10.1371/journal.pone.0162161>.
- Scherer, K.R., 1979. Non-linguistic indicators of emotion and psychopathology. In: Izard, C.E. (Ed.), *Emotions in personality and psychopathology*. Plenum, New York, pp. 495–529.
- Scherer, K.R., 1995. Expression of emotion in voice and music. *J. Voice* 9, 235–248.
- Shapiro, N.T., Hippe, D.S., Ramírez, N.F., 2021. How chatty are daddies? An exploratory study of infants' language environments. *J. Speech, Lang., Hear. Res.* 64 (8), 3242–3252.
- Siniscalchi, M., d'Ingeo, S., Minunno, M., Quaranta, A., 2018. Communication in dogs. *Animals* 8 (8). <https://doi.org/10.3390/ani8080131>.
- Skiendziel, T., Röscher, A.G., Schultheiss, O.C., 2019. Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System Scoring. *PLoS ONE* 14 (10), e0223905.
- Smith, S.L., 1983. Interactions between pet dogs and family members: an ethological study. In: Katcher, A.H., Beck, A.M. (Eds.), *New Perspectives on Our Lives with Companion Animals*. University of Pennsylvania Press, Philadelphia, pp. 29–36.
- Soderstrom, M., 2007. Beyond babytalk: re-evaluating the nature and content of speech input to preverbal infants. *Dev. Rev.* 27 (4), 501–532.
- Song, J.Y., Demuth, K., Morgan, J., 2010. Effects of the acoustic properties of infant-directed speech on infant word recognition. *J. Acoust. Soc. Am.* 128 (1), 389–400.
- Suppnen, E., Huotilainen, M., Ylinen, S., 2019. Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behav. Dev.* 57 (May), 101346 <https://doi.org/10.1016/j.infbeh.2019.101346>.
- Tamis-LeMonda, C.S., Kuchirko, Y., Luo, R., Escobar, K., Bornstein, M.H., 2017. Power in methods: language to infants in structured and naturalistic contexts. *Dev. Sci.* 20 (6), e12456.
- Tannen, D., 2004. Talking the dog: framing pets as interactional resources in family discourse. *Res. Lang. Soc. Interact.* 37 (4), 399–420.
- Thiessen, E.D., Hill, E.A., Saffran, J.R., 2005. Infant-directed speech facilitates word segmentation. *Infancy* 7 (1), 53–71.
- Topál, J., Miklósi, Á., Csányi, V., Dóka, A., 1998. Attachment behavior in dogs (*Canis familiaris*): a new application of Ainsworth's (1969) strange situation test. *J. Comp. Psychol.* 112, 219–229.
- Trainor, L.J., Austin, C.M., Desjardins, N., 2000. Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychol. Sci.* 11 (3), 188–195. <https://doi.org/10.1111/1467-9280.00240>.
- Trainor, L.J., Desjardins, R.N., 2002. Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychon. Bull. Rev.* 9 (2), 335–340.
- Venables, W.N., & Ripley, B.D. (2002). *Modern Applied Statistics with S* (4th ed.). Springer. (<https://www.stats.ox.ac.uk/pub/MASS4/>).
- Wang, Y., Houston, D.M., Seidl, A., Frühholz, S., Belin, P., 2018. Acoustic properties of infant-directed speech. *Oxf. Handb. Voice Percept.* 1, 93–116.
- Weinstein, N., Baldwin, D., 2022. Reification of infant-directed speech? Exploring assumptions shaping infant-directed speech research. *Cult. Psychol.* 1354067×221147683.
- Werker, J.F., McLeod, P.J., 1989. Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol. / Rev. Can. De. Psychol.* 43 (2), 230–246. <https://doi.org/10.1037/h0084224>.
- Werker, J.F., Pegg, J.E., McLeod, P.J., 1994. A cross-language investigation of infant preference for infant-directed communication. *Infant Behav. Dev.* 17 (3), 323–333.