



OPEN

## Six facial prosodic expressions caregivers similarly display to infants and dogs

Anna Gergely<sup>1</sup>✉, Édua Koós-Hutás<sup>1,2</sup>, Lőrinc András Filep<sup>1</sup>, Anna Kis<sup>1</sup> & József Topál<sup>1</sup>

Parents tend to use a specific communication style, including specific facial expressions, when speaking to their preverbal infants which has important implications for children's healthy development. In the present study, we investigated these facial prosodic features of caregivers with a novel method that compares infant-, dog- and adult-directed communication. We identified three novel facial displays in addition to the already described three facial expressions (i.e. the '*prosodic faces*') that mothers and fathers are typically displaying when interacting with their 1–18 month-old infants and family dogs, but not when interacting with another adult. The so-called *Special Happy* expression proved to be the most frequent face type during infant- and dog-directed communication which always includes a Duchenne marker to convey an honest and intense happy emotion of the speaker. These results suggest that the '*prosodic faces*' play an important role in both adult-infant and human-dog interactions and fulfil specific functions: to call and maintain the partner's attention, to foster emotionally positive interactions, and to strengthen social bonds. Our study highlights the relevance of future comparative studies on facial prosody and its potential contribution to healthy emotional and cognitive development of infants.

**Importance of infant-directed prosody.** When interacting with infants, adults adjust their communication style to the needs and abilities of the baby and this adjustment process can be characterized by specific intonation, uttering, body, and facial movements (for a review see<sup>1</sup>). Caregivers use this specific multimodal infant-directed (ID) prosody for multiple although not mutually exclusive purposes. These functions can be divided into two main categories: (i) cognitive functions (language tutoring—e.g.<sup>2</sup>; conveying information about the talker's intentions and identity—e.g.<sup>3,4</sup>) and (ii) affective functions (getting and maintaining the infant's attention—e.g.<sup>5,6</sup>; enhancing positive interactions with the infant—e.g.<sup>7</sup>; strengthening bonding between the interaction partners—e.g.<sup>8</sup>; expressing emotions—e.g.<sup>9</sup>; supporting emotional synchrony between interactants—e.g.<sup>10</sup>).

Several studies provided evidence that effective prosody, which can accomplish the aforementioned cognitive and affective functions, contributes greatly to young children's cognitive and emotional development. Less pronounced ID prosody is one of the typical symptoms of postpartum depression (PPD), and it is a potential risk factor for poor infant mental health (e.g.<sup>11</sup>). In line with this, children of mothers with PPD are more vulnerable to developing mental disorders later in life (for a review see<sup>12</sup>). It has been also shown that less exaggerated and less salient prosody of depressed mothers leads to failure in simple associative tasks in 4-month-old infants<sup>13</sup>. Moreover, infants of healthy mothers, who have been trained how to produce typical ID prosody, were more successful in linguistic tasks than their contemporaries whose mothers did not go through such training<sup>14</sup>.

While ID prosodic features are clearly multimodal exchanges, research has so far mainly focused on the acoustic components, and less is known about the visual domain, especially the specific facial expressions, that adults display when interacting face-to-face with infants. Without detailed investigations of ID facial prosody, we have limited knowledge of how and to what extent the visual domain contributes to healthy development of infants.

**Infant-directed facial prosody.** Chong and co-workers' paper<sup>15</sup> is one of the few studies focusing on 'visual prosody' in mother-infant interactions. The authors in this study have identified and described unique muscle movements of mothers in the form of three distinct facial expressions, each conveying a mixture of positive emotions like joy, happiness, surprise, excitement, and interest. These facial expressions can also be observed

<sup>1</sup>Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, 2 Magyar Tudósok Krt., Budapest, Hungary. <sup>2</sup>Doctoral School of Psychology, Eötvös Loránd University, 46 Izabella U., Budapest, Hungary. ✉email: gergely.anna@ttk.hu

in adult-directed (AD) communication, yet differ from them in certain characteristics and/or intensity<sup>16</sup>. One typical infant-directed (ID) facial expression is the so-called ‘mock surprise’ which was described by Stern<sup>17</sup>. It is characterized by exaggerated brow raise, open-stretched mouth with a hint of a smile, and accompanied frequently with ‘Ooooo’ vocalizations<sup>15</sup>. Another ID-specific facial display is the ‘fish face’<sup>17</sup> characterized by puckered lips with a moderate smile and brow raise. Authors suggested a soothing-comforting function of this facial expression<sup>15</sup>. Third, and the most frequent one is the so-called *Special Happy* expression, characterized by intense lip corner pull, cheek raise, and slightly open mouth<sup>15</sup>. It’s important to note, that according to Chong et al.<sup>15</sup> all three ID faces contain some type of smile, and the *Special Happy* is basically an intense Duchenne smile (in which zygomatic activity is combined with contraction of the orbicularis oculi—see<sup>18</sup>). Since all ID-specific expressions are typically displayed both by English and Chinese mothers, Chong et al.<sup>15</sup> concluded that these are probably universal (culture independent) and highlighted the importance of further investigations of the facial expressions that potentially convey an extra layer of meaning in infant-directed interactions.

Another interesting aspect of ID communication is the effect of the speaker’s gender on facial prosody. Although compared with mothers, fathers are underrepresented in the literature on adult-to-infant communication, some studies have provided evidence of gender-based differences in the use of facial prosody. It has been shown, for example, that distinct smile types (i.e. basic- or Duchenne smiles) can occur when observing mother-infant as opposed to father-infant dyads during natural playing sessions<sup>19</sup>. There is also evidence that parents have different interactive styles and father-infant interactions are more playful and arousing with sudden peaks in valence and different temporal patterns of happy emotion when compared with mother-infant interactions<sup>10,20</sup>. Because of differences in maternal and paternal face-to-face interactions with infants and due to the importance of fathers in healthy development of infants<sup>21</sup>, gender differences in facial prosody call for a more detailed investigation.

In addition to the gender-specific features, context-related (i.e. situation specific) differences in ID facial prosody is also worthy of mention. Facial prosody is strongly affected by the situation (e.g. naturalistic versus laboratory observations), as well as the linguistic and affective content of the interaction (see, e.g.<sup>22</sup> for a review). Mothers smile more when they sing than when they talk to their infants<sup>23</sup>. Mothers tend to produce more Duchenne smiles during an object play situation (in which they were speaking freely) as opposed to book reading or singing (with fixed sentences), while in fathers, book reading elicits more Duchenne smiles than other play situations<sup>19</sup>. Chong and co-workers<sup>15</sup> recorded mothers’ ID and AD facial expressions only during a free speech situation (i.e. without any instruction or restriction to the linguistic content) in which they were instructed to talk about six topics for at least 15 s. These topics were aimed to evoke different emotional content and expressions, like “tell a story about the first diaper change to your baby” or “tell a story about your feelings toward your baby” etc. Importantly, however, the authors did not analyse whether or not ID facial prosody of the mothers differed in terms of topics, thus the situation specific features of ID specific facial expressions (i.e. *Special Happy*, *Mock Surprise* and *Fish Face*) remained unknown.

**A comparative framework for studying facial prosody.** It is known, however, that ID prosody may not be uniquely used toward pre-verbal humans; similar speech forms have also been described toward pet companions such as dogs<sup>24,25</sup>, parrots<sup>26</sup> and cats<sup>2,27</sup>. This potentially suggests that ID-like communication forms are widespread towards non-speaking con- and heterospecific partners. Based on the increasingly important role of dogs in interspecific comparative investigations<sup>28,29</sup> here we propose that human–dog interaction in general, and dog-directed (DD) prosody in more particular, is a feasible option for studying the role that facial prosody plays in ID communication.

There is growing evidence supporting the viability of this approach. Family dogs show infant-like sensitivity both to the acoustic and to the facial features of ID prosody, like heightened pitch and happy facial expression<sup>30–33</sup>. This suggests that the attention getting function of ID and DD prosody might work similarly. Dogs also show similar patterns of attachment behaviour toward their caregiver as infants toward their mothers<sup>34</sup>. It is reasonable to assume, therefore, that DD prosody has an important role in enhancing the human–dog bond. Similarly to ID communication, gender- and situation specific features are also present in dog–human interactions<sup>25,35</sup>, which also make this comparative approach feasible and productive.

Besides these similarities, there are important differences between human infants’ and family dogs’ cognitive abilities and emotional processing. Dogs are non-conspecific social partners with different anatomy and the capacity to display facial expressions as compared to humans<sup>36,37</sup>. There is also evidence that they process human faces differently than humans do (e.g.<sup>38,39</sup>). These differences can have a significant effect on the reciprocity of emotional expressions, a key function of prosody, and suggest that mirroring events (i.e. when mothers and infants display similar facial expression one after another<sup>40,41</sup>) are probably rare or completely missing from an interaction with a dog. Last but not least, dogs, compared to infants, have a limited understanding of word meaning<sup>42,43</sup> and intentions of others<sup>44</sup>, and are not capable of forming higher-order representations<sup>45</sup>. In line with these findings, Burnham and co-workers<sup>2</sup> showed that hyperarticulation of vowels is missing from DD as opposed to ID communication. They suggest that hyperarticulation has a didactic function (language tutoring) which is presented only when we are talking to prospective speakers (i.e. infants) but not when speaking to non-human animals (see also<sup>25,26</sup>). Thus it is reasonable to assume that unlike in ID communication, cognitive functions of prosody have a minor, if any, role in DD communication. In sum, these similarities and differences make DD prosody a feasible model to study the species specific features as well as separate cognitive and affective functional features of facial prosody.

**Aims and hypothesis.** In light of the above research findings, the present paper aims to provide a detailed analysis of maternal and paternal facial prosodic characteristics of ID communication in comparison to AD

and DD communication during natural face-to-face interactions. In addition, we were also curious about the gender- and situation specific features of ID, DD, and AD facial prosody. To do so, we recorded the speaker's facial prosody during two free-speech situations with different object-related playing tasks and one-fixed speech situation (see Methods for details). Based on previous studies, these situations have the potential to evoke typical ID, DD and AD communication style<sup>25</sup> and are also feasible to study gender-related features of facial prosody toward infants<sup>19,22,23</sup>. We coded and analysed the occurrence (i.e. frequency) and intensity of facial expressions displayed during ID, DD and AD communication including those three ones that have been described in earlier studies (i.e. *Special Happy*, *Mock Surprise*, and *Fish Face*<sup>15,17</sup>). The effects of the speaker's gender, as well as the situation specific details of the facial expressions were also investigated by comparing the three situations within and between female and male speakers.

The questions we set out to address were the following:

(i) Whether speakers display similar facial prosody toward infants, dogs, and adults? Based on the aforementioned similarities and differences between communication directed towards infants' and dogs<sup>2,24–26</sup> we hypothesised an increased, but not identical, presence of prosodic facial expressions in both infant- and dog- directed (as compared to adult-directed) communication.

(ii) Whether facial prosody toward infants, dogs, and adults differ between female and male speakers and/or situations? Since previous studies suggested that happy facial expressions, including different smile types, differ between situations and this difference was affected by the speaker's gender<sup>19</sup>, we hypothesized interaction effects between gender and situation in terms of types (and frequency) of facial prosodic features. More specifically, more frequent and intense *Special Happy* expression, that includes a Duchenne smile, are expected during object-related situations as opposed to the fixed-speech situation in female speakers, and an opposite tendency is predicted in male speakers<sup>19</sup>. We also hypothesized that the occurrence and intensity of the facial prosodic features will be different in the more naturalistic free speech and less naturalistic fixed speech situations in both genders<sup>22</sup>.

## Results

**Categorization of facial expressions: the 'prosodic faces'.** We have found that 65.2% of the pictures fit into the three predetermined categories of facial expressions (*Fish Face*, *Mock Surprise*, *Special Happy*—see<sup>15,17</sup>). But in contrast to former studies, we also provided a detailed description of facial movements (including relevant AUs, activation percentages, and intensities) using the Facial Action Coding System (FACS) (see Table 1):

**Fish face—FF.** Similarly to previous studies<sup>15,17</sup> we found that this facial expression can be characterized by the crinkled lips (activation of AU18, AU22) and the slightly open mouth create (AU25, see Fig. 1a).

**Mock surprise—MS.** Similarly to previous studies<sup>15,17</sup> we found that this facial expression can be characterized by the simultaneous appearance of the raised eyebrows (AU1 + 2) and wide-open eyes (AU5) and mouth (AU26, AU27, see Fig. 1b).

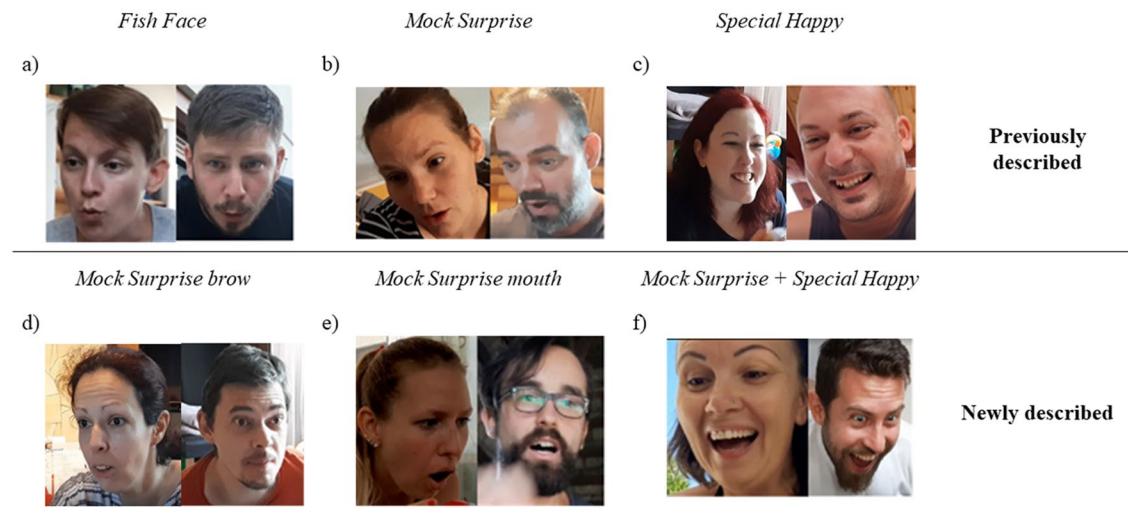
**Special Happy- SH:** In line with previous studies<sup>15,17</sup> we found that this facial expression can be characterized by a large lip corner pull (AU12) accompanied by a definite cheek raise (the Duchenne marker, corresponding to AU 6) and slightly opened mouth (AU25, Fig. 1c).

Importantly, however, we have also successfully identified three new facial expressions as a result of FACS coding (see Table 1):

	Special happy	Mock surprise	Fish face	Mock surprise brow	Mock surprise mouth	Mock surprise + special happy
N (1056 in total)	543	114	32	180	72	115
Brow (AU1 + 2)	<b>0.09 ± 0.41</b> 12.38%	<b>3.07 ± 1.73</b> 93.86%	<b>0.94 ± 1.7</b> 56.25%	<b>3.03 ± 1.35</b> 100%	<b>0.58 ± 0.9</b> 44.44%	<b>3.95 ± 1.27</b> 100%
Eyelid (AU5-AU44)	<b>-0.83 ± 1.38</b> 52.3%	<b>1.78 ± 1.71</b> 69.64%	<b>0.5 ± 1.29</b> 42.86%	<b>1.62 ± 1.59</b> 63.84%	<b>0.5 ± 1.22</b> 35.71%	<b>1.73 ± 1.95</b> 91.96%
Cheek (AU6)	<b>4.35 ± 0.84</b> 100%	<b>0.09 ± 0.39</b> 6.14%	<b>0.61 ± 0.99</b> 35.48%	<b>0.12 ± 0.44</b> 8.33%	<b>0.04 ± 0.26</b> 2.77%	<b>4.02 ± 1.16</b> 100%
Lip corner (AU12)	<b>4.27 ± 0.86</b> 100%	<b>0.79 ± 1.03</b> 48.25%	<b>0.06 ± 0.35</b> 3.13%	<b>0.94 ± 1.06</b> 55%	<b>0.5 ± 0.95</b> 27.78%	<b>3.93 ± 1.06</b> 99.13%
Lip pucker (AU18)	<b>0.05 ± 0.27</b> 3.5%	<b>0.9 ± 1.57</b> 35.09%	<b>3.25 ± 1.08</b> 100%	<b>0.55 ± 1.05</b> 28.89%	<b>0.42 ± 1.11</b> 16.67%	<b>0.35 ± 0.91</b> 16.52%
Lips part (AU25)	<b>4.32 ± 1.11</b> 98.12%	<b>4.42 ± 1.25</b> 97.35%	<b>2.97 ± 1.62</b> 93.55%	<b>3.15 ± 1.84</b> 86.36%	<b>4.81 ± 0.62</b> 100%	<b>4.3 ± 1.01</b> 100%
Mouth opening-jaw drop (AU26-27)	<b>0.32 ± 0.63</b> 33.77%	<b>1.96 ± 1.18</b> 92%	<b>0.81 ± 0.65</b> 79.31%	<b>0.42 ± 0.58</b> 47.9%	<b>2.37 ± 1.09</b> 100%	<b>1.31 ± 1.55</b> 62.83%

**Table 1.** Intensity of the facial muscle movements presented as separated Action Units (AUs, mean ± SE) and percentage (%) of AU activation (i.e. intensity score differed from 0) in the six 'prosodic faces'. Bold highlight represents relevant AUs (i.e. which was active in more than 50% of the cases).

## The ‘prosodic faces’



**Figure 1.** The ‘prosodic faces’: three facial displays from Stern<sup>17</sup> and Chong et al.<sup>15</sup> and three additional faces that we have identified in the present study.

**Mock surprise brow (MS(brow)).** The MS(brow) lacks great movements of the mouth, but a pronounced raise of the eyebrows are present similarly to MS (AU1 + 2, Fig. 1d).

**Mock surprise mouth (MS(mouth)).** The typical feature belonging to this expression is the wide-open mouth (AU26, AU27) occurring in MS, but the upper side of the face stays rather neutral, and no pronounced eyebrow movements are observable (Fig. 1e).

**Mock surprise + special happy (MSSH).** This label refers to a particular combination of the MS and SH expressions characterized by eyebrow raise (AU1 + 2) as well as cheek raise (AU6) and an intensive lip corner pull (AU12, Fig. 1f).

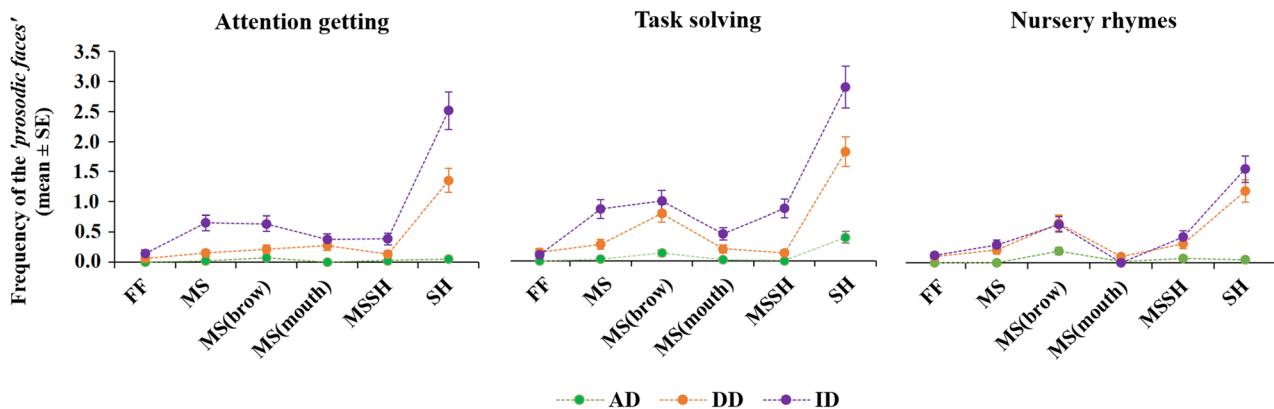
Although all of these ‘novel’ expressions are mixtures or subtypes of the above-mentioned predetermined facial expressions (FF; MS; SH), these faces have characteristics clearly distinguishable from those of described by Chong et al.<sup>15</sup> and Stern<sup>17</sup>. Therefore we decided to consider and analyse these six facial expressions (hereinafter called ‘prosodic faces’) separately.

**Frequency analysis.** As a next step we analysed the effects of partner, situation, speaker’s gender and the type of the expression on the frequencies of the ‘prosodic faces’. After removing the non-significant interaction terms, the final model revealed a significant three-way interaction effect of Partner × Situation × Face Type ( $F_{34,2208} = 1.78, p = 0.004$ ) and two-way interaction effect of Gender × Face Type ( $F_{5,2208} = 3.48, p = 0.006$ , for details see Supplementary Material 1.1 and Table S1a-c).

We found that almost all of the prosodic face types were more frequently expressed during ID and DD as opposed to AD communication during the ‘Task solving’ situation (pairwise comparisons  $p < 0.05$ , Table S1a, Fig. 2). The only exception was the Fish Face (FF) which showed only marginal significance between ID and AD communication during this situation ( $p = 0.05$ , for details see Table S1a, Fig. 2). During ‘Attention getting’ situation, all face types were more frequent during ID than in AD communication, while only MS, MS(mouth) and SH faces were more frequent in DD as opposed to AD communication (all pairwise  $p < 0.05$ ) and MS(brow) showed marginal significance ( $p = 0.05$ , Table S1a, Fig. 2). In comparison to AD communication, during ‘Nursery rhymes’ situation, the MS, MS(brow), MSSH and SH faces were more frequent during ID and DD communication (all pairwise  $p < 0.05$ , Table S1a, Fig. 2).

Results also showed that majority of the ‘prosodic faces’ were more frequent toward infants than toward dogs during the free speech situations (i.e. ‘Attention getting’ and ‘Task solving’: MS, MSSH and SH; all pairwise  $p < 0.05$ , in ‘Attention getting’ MS(brow)  $p = 0.005$ , in ‘Task solving’ MS(mouth) 0.028; Table S1a, Fig. 2). At the same time, no such difference has been found between DD and ID communication during the fixed speech situation (i.e. ‘Nursery rhymes’, all pairwise  $p > 0.05$ , Table S1a, Fig. 2).

Moreover, frequency analysis revealed a significant difference between face types, as Special Happy (SH) proved to be the most frequently displayed facial expression in ID and DD communication during the free speech situations (i.e. all pairwise  $p \leq 0.001$  in ‘Attention getting’ and ‘Task solving’; Table S1b, Fig. 2). MS(brow) proved to be the second most frequent face type and it was displayed more commonly than FF or MS(mouth) in DD and ID communication during ‘Task solving’ and ‘Nursery rhymes’ situations (all pairwise  $p < 0.02$ , Table S1b).



**Figure 2.** Mean (+SE) frequencies of the '*prosodic faces*' toward the three partners and situations. AD adult-directed, DD dog-directed, ID infant-directed. FF fish face, MS mock surprise, MSSH mock surprise + special happy, SH special happy.

Interestingly, we found that *Fish Face* (FF) was rarely displayed irrespectively of the partner and situation (for exact p-values see Table S1b, Fig. 2).

Our results showed that situation had an effect on the facial prosodic features of ID, DD and AD communication. Speakers displayed SH expression in AD communication, MS(brow) expression in DD- and MSSH expression in ID communication more during 'Task solving' situation compared to 'Attention getting' (all pairwise  $p < 0.02$ , Table S1c, Fig. 2). Similarly, SH toward all three partners and MS and MSSH expression in ID communication have been displayed more frequently during 'Task solving' situation compared to the 'Nursery rhymes' (all pairwise  $p < 0.05$ , Table S1c, Fig. 2). Interestingly, the frequency of MS(mouth) showed a different tendency in ID communication, as it was more common during 'Nursery rhymes' as opposed to 'Task solving' ( $p < 0.001$ , Table S1c, Fig. 2). Comparison of frequencies between 'Attention getting' and 'Nursery rhymes' situations showed a mixed picture. MS and SH expressions were more common in ID communication during 'Attention getting', while MS(brow) in DD and MS(mouth) in ID communication were more frequent during 'Nursery rhymes' (all pairwise  $p < 0.03$ , Table S1c, Fig. 2).

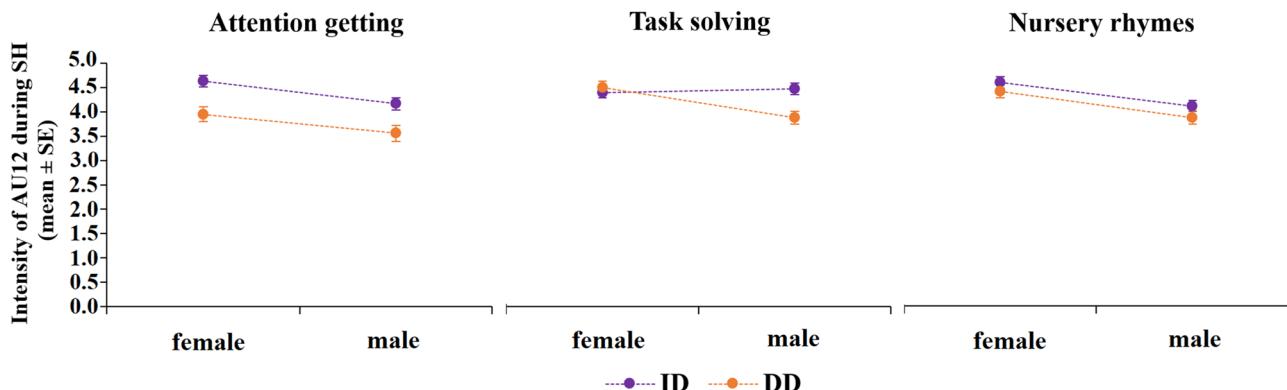
According to the Gender  $\times$  Face Type interaction, pairwise comparisons revealed similar frequencies in all six '*prosodic faces*' in female and male speakers (all pairwise  $p > 0.3$ , Table S2). Both genders displayed SH and MS(brow) expressions more frequently than the FF face type (all pairwise  $p < 0.01$ , Table S2, Figure S1). However, male speakers expressed the mouth and brow variation of *Mock Surprise* (MS(brow)) and *MS(mouth)* less frequently than the *Special Happy*, while no such difference could be found in female speakers (see Table S2, Figure S1).

**Intensity score analysis.** We also examined the intensity of the movement of facial muscles (AUs) involved in the formation of the '*prosodic faces*' during ID and DD communication. Models revealed several intensity differences between ID and DD facial prosody as well as between situations (see Supplementary Material 1.2 and Table S3 for full details).

Most common tendencies of the intensity results were the following:

- (1) ID and DD facial prosody differed in AU intensity in 10 out of the total of 25 models, and all 10 revealed greater intensity of the facial muscle movements toward infants than toward dogs (in all models  $F > 3.0$ ,  $p < 0.05$ , see Table S3). Is important to note that in the most frequent SH face type, all relevant AUs were moving more intensively during ID, than in DD communication.
- (2) In 10 out of the 14 significant models on AU intensity, female and male speakers were displaying similarly intense facial movements. If gender of the speaker had an effect on AU intensity, it typically showed more intense facial movements in female than in male speakers. That is, more intense lip corner pulling (AU12) and lips part (AU25) movements in *Special Happy*, more intense eyebrow raise (AU1+2) in MS(brow), and more intense mouth opening during 'Task solving' situation in *Fish Face* expression (all  $F > 7.8$ ,  $p \leq 0.006$ , Table S3). One exception was the eyelid movements (AU5/44) during *Special Happy* expression, which was more intense in male than in female speakers ( $F_{1,495} = 4.88$ ,  $p = 0.028$ ).
- (3) Situation had varying effects on AU intensity and these effects often manifested themselves through an interaction with the type of the partner and/or speaker's gender (see Table S3). In five out of seven models that revealed significant effect of situation, pairwise comparisons showed more intense AU movements during 'Task solving' and/or 'Nursery rhymes' as opposed to 'Attention getting' (in all 5 models  $p < 0.05$ , Table S3).

Figure 3 demonstrates an example of the results of the intensity analyses: the effects of gender and partner on the intensity of lip corner pulling muscle (AU12, a typical component of *Special Happy* expression) in the different task situations (for more details see Supplementary Material 1.2 and Table S3).



**Figure 3.** Mean (+ SE) scores of intensity of the lip corner puller (AU12) of male and female speaker's *Special Happy* (SH) face in the three experimental situations. DD dog-directed, ID infant-directed speech.

## Discussion

The aim of the present study was to provide an AU-based description and comparative analysis of maternal and paternal facial prosodic features that accompany infant- dog- and adult-directed communication. Occurrence (i.e. frequency) and intensity of facial prosody were examined during natural face-to-face interactions and analysed in relation to the given situation and the speaker's gender.

Our study provides evidence for the first time that infant-directed (ID) and dog-directed (DD) communication can be characterized by similar specific facial prosodic expressions of female and male speakers. In addition to the previously described three 'prosodic faces' toward infants<sup>15,17</sup>, here we have identified three novel expressions. Our results indicate that all six facial prosodic expressions are typical in both ID and DD communication but rare or completely missing form adult-directed (AD) communication. Importantly, we did not find any ID specific facial expression that was missing from DD prosody, note however, that the six 'prosodic faces' were less frequently and less intensively expressed by both male and female speakers toward dogs than towards infants. These results have important implications regarding the potential functions of ID facial prosody.

First, it is reasonable to assume that the '*prosodic faces*' are expressed for multiple purposes that are similarly important when communicating with a preverbal infant and a pet dog. Based on previous studies on the functional similarities and differences between infants' and dogs' cognitive and emotional abilities<sup>33,34,43,45</sup>, we may assume that the '*prosodic faces*' have an important role in calling and maintaining the partner's attention, in increasing social bond between the partners and in fostering emotionally positive interactions. In line with previous findings on acoustic prosody these results further support the notion that people tend to use similar cues to engage infants and dogs in communicative interactions and this assumption can be extended to the facial features of prosody<sup>24,25</sup>.

At the same time, we can assume that the significantly less intense and frequent facial prosody toward dogs than toward infants might due to functions that are important when interacting with an infant, but less relevant when interacting with a dog (i.e. language tutoring, expressing intentions and in facilitating emotional synchrony between the speaker and the partner). Note, however, that our results are not conclusive, thus further studies (comparative analyses of the speaker's and the partner's facial expressions in ID and DD communicative changes during language tutoring and intention expressing situations) are needed to verify these assumptions about emotional synchrony and mirroring events.

Our results also indicate that one ID- and DD-specific facial expression, the *Special Happy*, significantly outweighs the other five ones in terms of frequency. This facial display dominated both ID and DD communicative exchanges regardless of the given situation, while there were faces, like the *Fish Face* and the *Mock Surprise mouth* variation that were rarely expressed. Huge variation in the number of occurrences suggests that distinct facial displays might have different roles and they contribute to the aforementioned functions of prosody to varying degrees. The *Special Happy* expression is basically an intense Duchenne smile that is typically accompanied with opened mouth<sup>15,17</sup>. The Duchenne smile is described as a spontaneous, genuine expression of an intense happy emotion that communicate strong social affiliation and interest to be engaged in social interaction<sup>18</sup>. Accompanied with opened mouth, the Duchenne smile is perceived even more happy as opposed to the closed mouth version of the same smile<sup>46</sup>. These indicate that female and male speakers of the present study intended to communicate their intense and honest happy emotion and their willingness to engage in social interaction in all task situations in which they interacted with their own preverbal infants or family dogs. The preference to look at positive facial expressions of humans has been shown in dogs and infants<sup>32,33</sup>, moreover, infants are more attentive toward smiles when accompanied by open-mouth (e.g.<sup>47</sup>). There is no doubt that such a smile has the potential to enhance positive interactions with an infant<sup>7</sup>, and now we can extend this assumption to canine companions.

In contrast to our prior hypothesis, the frequency of the '*prosodic faces*' was not influenced by the interaction of the speaker's gender and the given situation<sup>19</sup>. At the same time, intensity analysis of the most frequent face type, the *Special Happy* showed such interaction effect during ID and DD communication. That is, male speakers, displayed a more intense cheek raise (Action Unit 6), an important component of the Duchenne smile<sup>18</sup>, during the object-play situation (i.e. 'Task solving') in comparison to the fixed speech situation (i.e. 'Nursery rhymes').

This is not true for female speakers: facial movements that are related to the Duchenne smile (i.e. Action Unit 6 & 12) were not influenced by the given situation. Note, however, that several Action Units showed more intense muscle movements in female than male speakers irrespective of the situation, including a more exaggerated lip corner pulling and mouth opening during *Special Happy* expression. Male speakers tended to open their eyes more than females, however, this difference was strongly dependent on the situation and the type of specific facial expression. These are in contradiction to previous findings on paternal and maternal smile types and emotional intensity of facial expressions toward infants<sup>19,20</sup>. This may be due to methodological differences between the present and previous studies, namely the age of the speakers, the age of the infant partners, differences in the experimental situations, and the measured variables, etc. In addition, cultural differences between American, Greek, and Hungarian parents can also have a significant effect on facial prosody (e.g.<sup>15,48,49</sup>), which raise interesting questions for future studies. In sum, the majority of our results, support the notion that facial displays accompanying infant-directed speech are rather universal among mothers and fathers (see<sup>50</sup>): paternal and maternal ID facial prosody share key features, and these similarities still exist in human–dog communicative exchanges.

In line with our hypothesis, we found several differences between free speech versus fixed speech situations. For example, during the fixed speech situation (i.e. ‘Nursery rhymes’) frequency of all ‘prosodic faces’ was similar during ID and DD communication, while several differences were found during the free speech situations (‘Task solving’ and ‘Attention getting’) in this regard. ‘Task solving’ situation evoked the most intense and frequent facial movements, including the most common *Special Happy* expression and this was so irrespective of the partner (i.e. infant/adult/dog). This finding is in line with our previous study on acoustic features of ID and DD prosody in which we observed the highest pitch<sup>25</sup> in a very similar ‘Task solving’ situation. Speakers in this situation were asked to hide a toy and encouraged their partners (infant, dog, or adult) to find it. Given that praise occurred typically in this ‘Task solving’ situation (when the partner was successful in finding the hidden toy) it is reasonable to assume a functional link between the ‘prosodic faces’ and praise. This further confirms Chong et al.’s<sup>15</sup> notion that the *Special happy* facial expression of mothers toward their infants is associated with praise.

Attention getting is usually listed among the main functions of prosody (e.g.<sup>6</sup>). Surprisingly, however, we found a lower rate of occurrence of the ‘prosodic faces’ in the ‘Attention getting’ situation than in the ‘Task solving’. Speakers in this situation were asked to attract their partners’ attention toward an object (i.e. a toy) as much as possible. Intense emotional facial expressions, however, can primarily serve to direct partners’ attention toward the face of the speaker (e.g.<sup>51</sup>). To effectively keep partners focused on the object, speakers thus may reduce their facial movements in the ‘Attention getting’ situation. However, more studies (comparisons of attention getting situations with and without the involvement of objects) are needed to confirm this conclusion.

One key finding of the current study is the identification and detailed characterization of three novel ID and DD specific facial expressions, in addition to the already published three faces by Chong and co-workers<sup>15</sup>. It is important to note that all newly identified ‘prosodic faces’ are variations and combinations of the previously described ones<sup>15</sup> but they were reliably distinguishable from them. The reason why earlier studies described only three types of facial expressions may stem from methodological differences in the studies. Namely, Chong and co-workers<sup>15</sup> measured the facial movements of mothers during one storytelling situation in which stories were carefully chosen to elicit a variety of emotions and related facial expressions (e.g., worried, happy) from the mother. In contrast, in the present study, we conducted three different situations that mimic everyday situations between the caregiver and the child. In two out of the three situations, speakers were allowed to talk freely, therefore the emotional content of their speech was not directly controlled by the experimental procedure. As facial expressions are highly influenced by the speaker’s intentions and emotions (e.g.<sup>52</sup>) it is possible that the emotionally less controlled situations evoked these novel and distinct faces. An alternative explanation would be that Chong et al.<sup>15</sup> also observed these facial expressions, but classified them all as *Mock Surprise* or *Special Happy* varying only in intensity. For instance, it is possible that *Mock Surprise brow* facial expression was considered as the slight variation of *Mock Surprise* that lacks a jaw drop but includes intense brow raising movement and thus is not worthy of a separate rating. However, Chong et al.<sup>15</sup> did not report a detailed analysis of AU intensity for the different facial expressions, so this possibility remains speculative. It should also be noted that differences in the infants’ age between the Chong et al.’s<sup>15</sup> and our study might serve as another plausible explanation for the aforementioned differences. In the present study, infants’ mean age was 10 months ( $\pm 4.1$  months), while in Chong et al. they were between 4 and 7 months<sup>15</sup>. As caregivers adjust their ID prosody to their infant’s capabilities and needs<sup>53</sup>, it is possible that these novel faces (i.e. *Mock surprise brow* and *mouth* variation as well as the *Mock Surprise + Special Happy* expression) are typically displayed when interacting with older, but not younger infants. Further studies are needed to investigate these possibilities.

In sum, the main findings of the current work can be summarized in the following way: (i) We provided evidence for the first time, that caregivers tend to use similar facial expressions when interacting with their preverbal infants and family dogs. (ii) Both female and male speakers used six distinct facial expressions, three were identified by the current study, that was typical when speaking to infants and dogs, but were rare or completely missing when speaking to adults. (iii) Facial prosody toward infants, dogs, and adults had context specific features, and situations that contain praise had the potential to evoke frequent and intense facial prosody. (iv) Our results highlighted the importance of infant-directed facial prosody in getting infant attention, conveying positive emotions through the interaction, and in strengthening the social bond between caregivers and infants. These results support the idea that the comparative investigation of facial prosody in human–dog and adult–infant communicative exchanges has the potential to broaden our knowledge about the significance of visual prosody and raises important questions for future research.

**Limitations of the study.** This comparative study can be considered as the first step toward uncovering the role facial prosody plays in adult–infant and human–dog social interactions. Therefore, it has several limita-

tions which should be investigated in future studies. (i) We studied ID, DD, and AD facial prosody during three situations that were evoking mostly positive emotions. We therefore could provide no information about visual prosody accompanied by negative emotions (e.g. fear, sadness, pain, etc.). (ii) Partner's behaviour (i.e. facial expressions, attention, etc.) was not recorded in the present study, therefore we did not provide any information about the potential mirroring event or emotion synchronization. (iii) We did not analyse acoustic prosody in the present study, thus dynamics of the two modules (acoustic and facial prosody) were not studied. (iv) It has been raised that a friendly but unfamiliar adult is not a feasible control for infant-directed prosody because the lack of attachment between the speaker and the adult partner<sup>54</sup>. Here we used an unfamiliar friendly adult as a partner during the adult-directed communication in order to produce comparable data with relevant former studies (e.g.<sup>15,19,25</sup>). (v) Last but not least, potential underlying factors, like mental state (including depression), personality, experience with infants and dogs etc. of the speakers, that can have an effect on facial prosody were not analysed either. Future comparative studies are needed to study the aforementioned questions, for which, our study provided several important results and a feasible comparative framework.

## Materials and methods

**Ethic statement.** This research was approved by the Human Research Ethics Committee (EPKEB) at Hungarian Academy of Sciences (No. 2015/23, 2022-85). In accordance with ethics approval, all parents completed an informed written consent to participate in the study and all methods were performed in accordance with the relevant guidelines and regulations of the EPKEB and the current laws of Hungary.

**Subjects.** Sample size of the current study was based on a power analysis considering the design and statistical tests to be used (G\*Power 3.1.9.7.). As there is no previous work that could provide specific effect size estimates, the a priori power analysis assumed a 'medium' effect size,  $\alpha = 0.05$  and power  $(1 - \beta) = 0.8$ . 42 participants from 22 families ( $N = 22$  females;  $N = 20$  males, age range: 18–45 years—for demographic details see Supplement Material 2.1 and Table S4) were recruited on a voluntary basis who had both a 1–18 months old child (13 girls, 9 boys, mean age 10.1 months  $\pm$  4.1 months SD) and an adult (over 1 year old) pet dog at home (16 females, 13 males, mean age  $\pm$  SD 4.9 years  $\pm$  2.9 years).

All participants spoke Hungarian as their native language and from all participant, informed consent has been obtained to publish the information/image(s) in an online open-access publication.

**Apparatus.** For video recordings we used a smartphone (Samsung Galaxy A70) on a tripod set to capture the participant's face. We used three separate toy balls in the ID (sponge ball with a smiling face), DD (rubber dog toy) and AD (plastic ball) interactions, same in size and colour, for the 'Attention getting' and 'Task solving' situations (for details see below). For the 'Nursery rhymes' situation participants were allowed to use the printed version of the rhymes if needed.

**Procedure.** Data acquisition took place in the participants' home. During the recordings a female experimenter (the 1<sup>st</sup> author of this paper) was present: she explained the tasks for the participants, controlled the starting and stopping of recordings and played the role of the partner during AD communication. Each participant was recorded speaking separately to their (i) infant (ID), (ii) dog (DD) and (iii) an adult experimenter (AD) in a within-subjects design. Video recordings of the participants were collected in two free speech situations ('Attention getting' and 'Task solving') in which they were encouraged to interact and speak freely without restrictions, and one fixed speech situation where they were asked to tell two nursery rhymes to the partner ('Nursery rhymes'). Each recording lasted 30 to 40 s/situation (average length was 117  $\pm$  11 s) and participants were encouraged to interact with all partners in a natural manner. Length of the recordings did not differ between ID, DD and AD communication (Linear GLMM,  $F_{2,123} = 0.744$ ,  $p = 0.477$ ). All three situations were recorded toward all three partners (ID, DD and AD). The order of ID, DD and AD communication were counterbalanced across participants.

'Attention getting': participants were instructed to call the partner's (infant, dog or adult) attention toward a small red ball, and they were asked to use the following words as frequently as possible: "red", "ball", "rolling" (in Hungarian these words are "piros", "labda", "gurul"). We asked this because these words in Hungarian contain all three vowels important for examining hyperarticulation (i, a, u,—see also<sup>25</sup>). We used three different balls for ID, AD and DD communication due to hygienic reasons (for details see Apparatus).

'Task solving': participants were instructed to play a hide-and-seek game with the partner by using the same red ball as in the 'Attention getting' situation. They were asked to hide the ball into one of their hands then ask the partner to find it. They were also instructed to use the words "red", "ball", "rolling" as frequently as possible during this phase too.

'Nursery rhymes': participants were asked to pick two rhymes they already know out of 8 and tell these rhymes to the partner.

**Data analysis.** As the first step, one of the experimenter (the 2nd author of this paper, Coder 1) watched all of the recordings (42 speakers  $\times$  3 partners  $\times$  3 situations = 378 videos in total) while aiming to identify all facial movements that differed from a neutral facial expression, including the three faces described by Stern and Chong et al.<sup>15,17</sup>. She, then saved each of these facial expressions as individual pictures which resulted in a total of 1084 pictures. It is important to note that we did not analyse duration of the facial expressions in the present paper. Some expressions, *Special Happy* for instance (see details below), are typically presented for long durations with varying intensity, but without changing into neutral or into other face types. In this case, it was included to our sample as one *Special Happy* expression, and the screen shot was taken at the most intense moment (as judged

by Coder 1). As a next step, Coder 1 classified and sorted the pictures which were fitting into the three categories according to Stern and Chong et al.<sup>15,17</sup>:

**Fish face (FF).** The name of this facial expression stems from the typical position which the crinkled lips and the slightly open mouth create. When examining the face's partition by Ekman et al.<sup>55</sup> the changes in the action unit (AU) 18 and 22 can be associated with this type of ID expression.

**Mock surprise (MS).** This term sums up the simultaneous appearance of the raised eyebrows and wide-open eyes and mouth with a lip corner pull (i.e. AU 12).

**Special happy (SH).** This particular expression is more than a casual smile because without exception the mouth is at least slightly open. Similar to the Duchenne smile, its most important features are the cheek raise (corresponding to AU 6) and lip corner pull (AU 12). According to Ekman et al.<sup>55</sup>, AU 12—i.e. facial changes produced by the zygomatic major muscles—are always an indicator of a positive emotional state. For differentiating it from a simple, but bright smile, the fine wrinkles around the eyes' outer side were taken into account.

Note, however, that 35.6% of the pictures did not fit neatly into any predetermined category. The experimenter therefore sorted these pictures into three additional categories on the basis of changes in facial muscle movements (*Mock Surprise brow; Mock Surprise mouth; Mock Surprise + Special Happy*—for a detailed description of these faces see Results). Inter-observer reliability scores were measured between two additional independent observers (the 1<sup>st</sup> and 3<sup>rd</sup> authors of this paper, Coder 2 and 3) for the observed facial expressions. Percentage agreements and Cohen's kappa coefficients were calculated for each category of facial expressions: % agreement ranged from 95.04 to 99.01 and Cohen's kappa values were between 0.84 and 0.91 showing a high level of reliability (Table 2). Note, however, that in case of disagreement between the three coders (Coder 1, 2 and 3), the categorization of the observed facial expression was based on majority opinion.

As a second step, Coder 2 and 3 also analysed the pictures using the golden-standard Facial Action Coding System (FACS<sup>56</sup>). This method defines the intensity of the given facial expressions by assigning values on a unique scale to specific areas of the face referred to as Action Units (AUs). 28 pictures have been excluded from FACS analysis because of the angle of the face or the light conditions made reliable coding impossible. This resulted a total of 1056 pictures for FACS coding. For detailed description of the analysed AUs see Supplementary Material (Table S5). In accordance with the instructions and criteria for calculating intensity levels illustrated in the *Manual for Facial Action Coding System*<sup>57</sup> Coder 2 and 3 assigned the corresponding intensity score to all movement types (0: *neutral/absent*; 1: *trace*; 2: *slight*; 3: *marked/pronounced*; 4: *severe*; 5: *maximum*). In case of the mouth opening and eyelid movements we used a slightly different scale (for details see Supplementary Material). Figure 4 shows one example of typical infant-directed facial expressions and corresponding FACS coding.

Inter-observer reliability scores were measured between two additional independent observers (Coder 2 and 3) for the FACS coding.

	SH	FF	MS	MSSH	MS (brow)	MS (mouth)
N (1084 in total)	547	33	118	110	203	73
%	98.72	99.01	96.61	98.19	95.42	98.81
Cohen's κ	0.84	0.86	0.84	0.91	0.86	0.91

**Table 2.** Percentage agreements and Cohen's kappa coefficients for the six facial expression types.



face part	Action unit	Intensity
eyebrows	AU(1+2)	5
eyelid	AU5	5
cheek	AU6	4
mouth opening	AU26	2
lip corner	AU12	4
lips part	AU25	5

**Figure 4.** A typical MSSH facial expression and the corresponding FACS coding.

Coder 2 (AG) coded 3 out of the 6 facial expression types (*SH*, *MSSH* and *FF*), while Coder 3 (LAF) coded the remaining 3 types (*MS*, *MS(brow)* and *MS(mouth)*). Inter-observer agreements for FACS coding were assessed by means of parallel coding of the 20% of the total sample. The integrated intra-class correlation coefficient (ICC) was 0.84, which is considered to be good<sup>58</sup>.

**Statistical analysis.** The effects of the partner (AD, DD & ID), situation ('Attention getting', 'Task solving', 'Nursery rhymes'), the speakers' gender (female/male) and their interactions on both the frequency and intensity of facial expressions were analysed by using Generalized Linear Mixed Models (GLMMs). Backward elimination technique was applied in all models in a stepwise manner for removing the non-significant interactions and main effects from the model. Identification number for each participants was included as a random grouping factor in order to control for repeated measurement. In all post-hoc comparisons Bonferroni correction was applied. For all statistical analysis we used IBM SPSS Statistics (version 23) software.

Frequency analyses were carried out by using GLMM with Negative Binomial regression. Fixed factors were Gender (female/male); Partner (AD, DD, ID); Situation ('Attention getting', 'Task solving', 'Nursery rhymes'), and Face Type (*SH*, *MSSH*, *MS*, *MS(brow)*, *MS(mouth)*, *FF*). All two- and three-way interactions were also included in the model (Gender × Partner; Gender × Face Type; Gender × Situation; Situation × Face Type; Situation × Partner, Partner × Face Type; Gender × Partner × Situation, Gender × Partner × Face Type, Gender × Situation × Face Type, Partner × Situation × Face Type).

For intensity analyses we used separate GLMMs for Gaussian distribution for each relevant AU in each face type (see Table S5 for the list of analysed AUs). We considered an AU relevant if it was activated (i.e. intensity score differed from 0) in that specific face type in at least 50% of cases (for details see Table 1).

Since all face types were markedly underrepresented in AD communication (only 48 out of 1084 faces) and the *FF* expression was completely missing, we decided to compare intensity between DD and ID communication only. In all models fixed factors were Gender (female, male); Partner (DD, ID) and Situation ('Attention getting', 'Task solving', 'Nursery rhymes'). All two- and three-way interactions were also examined (Partner × Situation; Partner × Gender; Gender × Situation; Gender × Partner × Situation).

## Data availability

The data analysed during the current study are available from the corresponding author on reasonable request. Raw data: frequency and intensity data can be found in Supplementary Dataset 1.

Received: 19 September 2022; Accepted: 21 December 2022

Published online: 17 January 2023

## References

1. Esteve-Gibert, N. & Guellai, B. Prosody in the auditory and visual domains: A developmental perspective. *Front. Psychol.* **9**, 1–10 (2018).
2. Burnham, D., Kitamura, C. & Vollmer-Conna, U. What's new, pussycat? On talking to babies and animals. *Science* **296**, 1435 (2002).
3. Fernald, A. *et al.* A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J. Child Lang.* **16**, 477–501 (1989).
4. Mehler, J., Bertoni, J., Barriere, M. & Jassik-Gerschenfeld, D. Infant recognition of mother's voice. *Perception* **7**, 491–497 (1978).
5. Fernald, A. Four-month-old infants prefer to listen to motherese. *Infant Behav. Dev.* **8**, 181–195 (1985).
6. Spinelli, M., Fasolo, M. & Mesman, J. Does prosody make the difference? A meta-analysis on relations between prosodic aspects of infant-directed speech and infant outcomes. *Dev. Rev.* **44**, 1–18 (2017).
7. Murray, L. & Trevarthen, C. Emotional regulation of interactions between two-month-olds and their mothers. in *Social Perception in Infancy* (eds. Field, T. & Fox, N.), 177–197 (Ablex Press, 1985).
8. Fernald, A. Meaningful melodies in mothers' speech to infants. in *Nonverbal Vocal Communication: Comparative and Developmental Approaches* (eds. Papousek, H., Jurgens, U. & Papousek, M.) (Cambridge University Press, 1992).
9. Scherer, K. R. Expression of emotion in voice and music. *J. Voice* **9**, 235–248 (1995).
10. Feldman, R. Infant-mother and infant-father synchrony: The coregulation of positive arousal. *Infant Ment. Health J.* **24**, 1–23 (2003).
11. Bettes, B. A. Maternal depression and motherese: Temporal and intonational features. *Child Dev.* **59**, 1089–1096 (1988).
12. Grace, S. L., Evindar, A. & Stewart, D. E. The effect of postpartum depression on child cognitive development and behavior: A review and critical analysis of the literature. *Arch. Womens. Ment. Health* **6**, 263–274 (2003).
13. Kaplan, P. S., Bachorowski, J. A., Smoski, M. J. & Hudenko, W. J. Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. *Psychol. Sci.* **13**, 268–271 (2002).
14. Morningstar, M., Garcia, D., Dirks, M. A. & Bagner, D. M. Changes in parental prosody mediate effect of parent-training intervention on infant language production. *J. Consult. Clin. Psychol.* **87**, 313–318 (2019).
15. Chong, S. C. F., Werker, J. F., Russell, J. & Carroll, J. Three facial expressions mothers direct to their infants. *Infant Child Dev.* **12**, 211–232 (2003).
16. Ekman, P. & Friesen, W. V. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues* (Englewood Cliffs, 1975).
17. Stern, D. N. Mother and infant at play: The dyadic interaction involving facial, vocal and gaze behaviors. in *The Effect of the Infant on its Caregiver* (ed. Lewis, M. Rosenblum, L.), 187–232 (Wiley, 1974).
18. Ekman, P., Davidson, R. & Friesen, W. V. The Duchenne smile: Emotional expression and brain physiology II. *J. Pers. Soc. Psychol.* **58**, 342–353 (1990).
19. Dickson, K. L., Walker, H. & Fogel, A. The relationship between smile type and play type during parent-infant play. *Dev. Psychol.* **33**, 925–933 (1997).
20. Kokkinaki, T. & Vasdekis, V. G. S. Comparing emotional coordination in early spontaneous mother-infant and father-infant interactions. *Eur. J. Dev. Psychol.* **12**, 69–84 (2015).
21. Lamb, M. E. *The Role of the Father in Child Development* (Wiley, 2004).
22. Fernández-Dols, J. M. & Crivelli, C. Emotion and expression: Naturalistic studies. *Emot. Rev.* **5**, 24–29 (2013).
23. Trehub, S. E., Plantinga, J. & Russo, F. A. Maternal vocal interactions with infants: Reciprocal visual influences. *Soc. Dev.* **25**, 665–683 (2016).
24. Hirsh-Pasek, K. & Treiman, R. Doggerel: Motherese in a new context. *J. Child Lang.* **9**, 229–237 (1982).

25. Gergely, A., Faragó, T., Galambos, Á. & Topál, J. Differential effects of speech situations on mothers' and fathers' infant-directed and dog-directed speech: An acoustic analysis. *Sci. Rep.* **7**, 1–10 (2017).
26. Xu, N., Burnham, D., Kitamura, C. & Vollmer-Conna, U. Vowel hyperarticulation in parrot-, dog- and infant-directed speech. *Anthrozoos A Multidiscip. J. Interact. People Anim.* **26**, 373–380 (2013).
27. de Mouzon, C., Gilbert, C., Di-Stasi, R. & Leboucher, G. How's my kitty? Acoustic parameters of cat-directed speech in human-cat interactions. *Behav. Processes* **203**, 104755 (2022).
28. Bunford, N., Andics, A., Kis, A., Miklósi, Á. & Gácsi, M. *Canis familiaris* as model for non-invasive comparative neuroscience. *Trends Neurosci.* **40**, 438–452 (2017).
29. Topál, J., Román, V. & Turcsán, B. The dog (*Canis familiaris*) as a translational model of autism: It is high time we move from promise to reality. *Wiley Interdiscip. Rev. Cogn. Sci.* **10**, 1–11 (2019).
30. Jeannin, S., Gilbert, C., Amy, M. & Leboucher, G. Pet-directed speech draws adult dogs' attention more efficiently than adult-directed speech. *Sci. Rep.* **7**, 1–9 (2017).
31. Trainor, L. J. & Zacharias, C. A. Infants prefer higher-pitched singing. *Infant Behav. Dev.* **21**, 799–805 (1998).
32. Kim, H. I. & Johnson, S. P. Do young infants prefer an infant-directed face or a happy face? *Int. J. Behav. Dev.* **37**, 125–130 (2013).
33. Müller, C. A., Schmitt, K., Barber, A. L. A. & Huber, L. Dogs can discriminate emotional expressions of human faces report. *Curr. Biol.* **25**, 1–5 (2015).
34. Topál, J., Miklósi, Á., Csányi, V. & Dóka, A. Attachment behavior in dogs (*Canis familiaris*): A new application of Ainsworth's (1969) strange situation test. *J. Comp. Psychol.* **112**, 219–229 (1998).
35. Prato-Previde, E., Fallani, G. & Valsecchi, P. Gender differences in owners interacting with pet dogs: An observational study. *Ethology* **112**, 64–73 (2006).
36. Bloom, T. & Friedman, H. Classifying dogs' (*Canis familiaris*) facial expressions from photographs. *Behav. Processes* **96**, 1–10 (2013).
37. Waller, B. M. *et al.* Paedomorphic facial expressions give dogs a selective advantage. *PLoS ONE* **8**, e82686 (2013).
38. Szabó, D. *et al.* On the face of it: No differential sensitivity to internal facial features in the dog brain. *Front. Behav. Neurosci.* **14**, 25 (2020).
39. Bunford, N. *et al.* Comparative brain imaging reveals analogous and divergent patterns of species and face sensitivity in humans and dogs. *J. Neurosci.* **40**, 8396–8408 (2020).
40. Lavelli, M. & Fogel, A. Interdyadic differences in early mother-infant face-to-face communication: Real-time dynamics and developmental pathways. *Dev. Psychol.* **49**, 2257–2271 (2013).
41. Markova, G. & Legerstee, M. Contingency, imitation, and affect sharing: Foundations of infants' social awareness. *Dev. Psychol.* **42**, 132–141 (2006).
42. Fukuzawa, M., Mills, D. S. & Cooper, J. J. More than just a word: Non-semantic command variables affect obedience in the domestic dog (*Canis familiaris*). *Appl. Anim. Behav. Sci.* **91**, 129–141 (2005).
43. Ramos, D. & Mills, D. S. Limitations in the learning of verbal content by dogs during the training of OBJECT and ACTION commands. *J. Vet. Behav.* **31**, 92–99 (2019).
44. Moore, R., Mueller, B., Kaminski, J. & Tomasello, M. Two-year-old children but not domestic dogs understand communicative intentions without language, gestures, or gaze. *Dev. Sci.* **18**, 232–242 (2015).
45. Udell, M. A. R., Dorey, N. R. & Wynne, C. D. L. Can your dog read your mind?: Understanding the causes of canine perspective taking. *Learn. Behav.* **39**, 289–302 (2011).
46. Messinger, D. S. Positive and negative: Infant facial expressions and emotions. *Curr. Dir. Psychol. Sci.* **11**, 1–6 (2002).
47. Walker-Andrews, A. S. Intermodal perception of expressive behaviors: Relation of eye and voice?. *Dev. Psychol.* **22**, 373–377 (1986).
48. Wörmann, V., Holodynski, M., Kärtner, J. & Keller, H. A cross-cultural comparison of the development of the social smile. A longitudinal study of maternal and infant imitation in 6- and 12-week-old infants. *Infant Behav. Dev.* **35**, 335–347 (2012).
49. Wörmann, V., Holodynski, M., Kärtner, J. & Keller, H. The emergence of social smiling: The interplay of maternal and infant imitation during the first three months in cross-cultural comparison. *J. Cross. Cult. Psychol.* **45**, 339–361 (2014).
50. Soderstrom, M. Beyond babble: Re-evaluating the nature and content of speech input to preverbal infants. *Dev. Rev.* **27**, 501–532 (2007).
51. Schindler, S. & Bublitzky, F. Attention and emotion: An integrative review of emotional face processing as a function of attention. *Cortex* **130**, 362–386 (2020).
52. Ekman, P. *Emotion in the Human Face* (Cambridge University Press, 1982).
53. Saint-Georges, C. *et al.* Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLoS ONE* **8**, 1–17 (2013).
54. Trainor, L. J., Austin, C. M. & Desjardins, N. Is infant-directed speech prosody a result of the vocal expression of emotion?. *Psychol. Sci.* **11**, 188–195 (2000).
55. Ekman, P., Friesen, W. V. & Ancoli, S. Facial Signs of Emotional Experience. *J. Pers. Soc. Psychol.* **39**, 1125–1134 (1980).
56. Ekman, P. & Friesen, W. V. *Investigator's Guide to the Facial Action Coding System, Part II* (Consulting Psychologists Press, 1978).
57. Ekman, P., Friesen, W. V. & Hager, J. *Facial Action Coding System. The Manual on CD ROM.* (2002).
58. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163 (2016).

## Acknowledgements

This research was supported by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund (K 128448, FK 142968, FK 128242), the Hungarian Brain Research Program (HBRP) 3.0 NAP and the János Bolyai Research Scholarship of the Hungarian Academy of Science. The authors are grateful to the parents and their families for participating in the experiment.

## Author contributions

Conceived and designed the experiment: A.G., J.T.; Performed the experiments: A.G.; Data analysis: A.G., É.K.H., L.F. Statistical analysis: A.G., É.K.H.; Wrote the paper: A.G., A.K., J.T. All authors gave final approval of the version to be published.

## Funding

Open access funding provided by ELKH Research Centre for Natural Sciences.

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-26981-7>.

**Correspondence** and requests for materials should be addressed to A.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023