Dog-directed speech: why do we use it and do dogs pay attention to it?

Article (Accepted Version)

Ben-Aderet, Tobey, Gallego-Abenza, Mario, Reby, David and Mathevon, Nicolas (2017) Dog-directed speech: why do we use it and do dogs pay attention to it? Proceedings of the Royal Society B: Biological Sciences, 284 (1846). ISSN 0962-8452

Dog-directed speech: why do we use it and do dogs pay attention to it?

Tobey Ben-Aderet[1], Mario Gallego-Abenza [2], David Reby[3] and Nicolas Mathevon[1,2]

[1]City University of New York, Hunter College, Department of Psychology, New York, USA

[2]University of Lyon/Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES/Neuro-PSI CNRS UMR9197, Saint-Etienne, France

[3]University of Sussex, School of Psychology, Falmer, United Kingdom

Pet-directed speech is strikingly similar to infant-directed speech, a peculiar speaking pattern with higher pitch and slower tempo known to engage infants' attention and promote language learning. Here we report the first investigation of potential factors modulating the use of dog-directed speech, as well as its immediate impact on dogs' behaviour. We recorded adult participants speaking in front of pictures of puppies, adult and old dogs, and analyzed the quality of their speech. We then performed playback experiments to assess dogs' reaction to dog-directed speech compared to normal speech. We found that human speakers used dog-directed speech with dogs of all ages and that the acoustic structure of dog-directed speech was mostly independent of dog age, except for sound pitch which was relatively higher when communicating with puppies. Playback demonstrated that, in the absence of other non-auditory cues, puppies were highly reactive to dog-directed speech, and that the pitch was a key factor modulating their behaviour, suggesting that this specific speech register has a functional value in young dogs. Conversely, older dogs did not react differentially to dog-directed speech compared to normal speech. The fact that speakers continue to use dog-directed with older dogs therefore suggests that this speech pattern may mainly be a spontaneous attempt to facilitate interactions with non-verbal listeners.

Keywords: Human-dog communication, pet-directed speech, infant-directed speech, hyperspeech, acoustic communication, dog cognition.

1. Introduction

When talking to their babies, human adults use a special speech register characterized by higher and more variable pitch, slower tempo, and clearer articulation of vowels than in speech addressed to adults [1-3]. This « infant-directed speech » has positive aspects in engaging and maintaining attention of babies and facilitating their social interactions with caregivers: infants as young as 7-weeks old show a preference for infant-directed speech over adult-directed speech [4]. Accordingly, infant-directed speech has been shown to increase cerebral activity more than adult-directed speech [5], meaning that infants are more engaged in what is being said to them when they listen to this special speech register. Infant-directed speech has also been hypothetized to facilitate language learning [6] by supporting the construction of phonetic and vowel categories [7,8], the clearer production of consonants [3], and the acquisition of new words [9]. This role in language learning is consistent with the decrease in the use and acoustic specificity of infant-directed speech that follows the development of language skills during the first year of the child [10-12]. At a proximal level, these dynamic changes could be explained by modifications of the baby's reactions to speech. As the baby grows up, he/she becomes more reactive to caregivers' solicitation and responds more specifically to meaningful sentences [13]. Promoting interaction thus becomes easier, which in return lessens the use of infant-directed speech. Another proximal explanation of the use of infant-directed speech could be that the morphological features of younger babies (large head, small nose and mouth = the « baby schema » described by Konrad Lorenz [14,15]) elicit infant-directed speech as part of caretaking behaviour. As these juvenile features become less prominent, their elicitation of infant-directed speech is expected to decrease. Thus, infant-directed speech appears to function as a communication signal that has evolved to accompany the cognitive development of babies and that may depend on proximate mechanisms that are both static (the « baby schema ») and dynamic (babies' attention response).

Dogs have been in close relationships with humans for thousands of years and this intimate proximity is reflected in many aspects of mutual understanding and empathy [16-21]. While more than eighty per cent of pet owners refer to themselves as "pet-parents" [22], adult women show similar brain activation patterns when presented with the picture of their dog and their own children [23]. Many dogs react to human vocal or gestural signals, and even feelings [20, 24]. Although dogs clearly do not possess the language ability, humans do change their speech patterns when talking to dogs using what is known as pet-directed speech, which shares similar structural properties with infant-directed speech (e.g. high-pitch register, slower tempo

[25, 26]).

Despite widespread interest in understanding the nature of the human-dog relationship, the proximate and ultimate factors that promote the use of pet-directed speech by human speakers remain unknown. The striking parallel between pet-directed speech and infant-directed speech may have different origins. Pet-directed speech may indeed constitute a spontaneous response of human speakers to juvenile characteristics shared by vertebrates' newborns (the "baby schema" hypothesis), or it may represent speakers' attempt at engaging an interaction with a non-verbal being (the "learning" hypothesis). The "baby schema" hypothesis predicts that humans should restrict the use of pet-directed speech to young puppies. In contrast, the "learning" hypothesis predicts that speakers should continue to use dog-directed speech with adult dogs as they do not develop the ability of language. Furthermore, the functional value of pet-directed-speech remains unknown, as, to our knowledge, the assumption that dogs respond more to pet-directed speech than to normal speech has not yet been tested.

The aim of the present study was thus to investigate whether the age of the dog receiver modulates the use and the properties of pet-directed speech. We then assessed the functional value of pet-directed speech by testing if it engages dogs' attention better than speech directed to human adults. To achieve this, we first recorded human speakers speaking in front of dogs' pictures and analyzed their vocal features. Second, we performed playback experiments on puppies and adult dogs to test their reaction to pet-directed speech versus to speech directed to human adults.

2. Material and methods

(i) Human speech recording and analysis

We selected 90 images of dogs' faces from the Internet with 30 dogs classified as "puppies" (< 1 year), 30 dogs classified as "adults" (1-8 years old) and 30 dogs classified as "old" (> 8 years), from a variety of dog breeds (the dogs' age and breeds were checked independently by two veterinarians; Supplementary Table 1). Each human speaker (n=30 women, aged 17-55) was then recorded (Zoom H4n digital recorder; sampling frequency = 44100 Hz) speaking in front of three of these pictures including one of a puppy, one of an adult dog, and one of an old dog (the pictures were presented using a smartpad). The set of three pictures differed between each recorded person. The images were successively presented to the recorded subject, in a balanced order between women (10 women were presented with the puppy first, 10 with the adult dog first and 10 with the old dog first). We also recorded the adult's voice in a control situation, without any dog picture, where the speaker was asked to speak to the researcher performing the recordings. This speech sequence was considered as human-directed speech. This control was obtained before the presentation of the set of dog pictures for 15 participants and after for the others. During each recording the adult repeated the same sentence, which was presented on the smartpad screen together with the dog's picture or in the absence of picture (control condition): "Hi! Hello cutie! Who's a good boy? Come here! Good boy! Yes! Come here sweetie pie! What a good boy!". For each participant we thus obtained a set of four recordings: "puppy-directed", "adult dog-directed", "old dog-directed" and "adult human-directed" (control) speech sequences of identical verbal content. Our recording procedure ensured that each speaker emitted exactly the same speech sequence in each recording condition. While recording the participants during an interaction with a real dog might have increased the ecological validity of our observations, the dynamic nature of the interaction would have inevitably led to variability in the uttered sentences, rendering the comparison between the acoustic features much more challenging.

Next, we performed acoustic analyses using PRAAT [27], and measured the following parameters (see Supplementary Methods): %voiced (percentage of the signal that is characterized by a detectable pitch), duration (total duration of the recording), mean F0, max F0, min F0 (respectively the mean, maximum and minimum fundamental frequency), F0CV (coefficient of variation of F0), inflex25 (minor intonation events), inflex2 (major intonation events), intCV (variability of the speech sequence's intensity), harm (harmonicity), jitter, shimmer, the first five formant frequencies of the speech sequence (F1, F2, F3, F4, F5).

(ii) Playback experiments to dogs

We performed playbacks to domestic dogs *Canis familiaris* to test (1) whether puppy-directed-speech is more effective than human-directed-speech in engaging a dog's attention, and if this effectiveness varies with dog's age, and (2) whether puppy-directed speech is more effective than adult dog-directed speech. The experiments were performed at the Bidawee animal shelter in Manhattan, New York (USA), between December 2015 and March 2016. The experimenter (TB) was volunteering in the shelter at the time of the study and spent several days a week with the participant dogs. All the tested dogs had a positive relationship with her prior to the tests. The experiments were conducted in a dedicated, spacious (3*4 m), room. All the tested dogs appeared comfortable in the testing situation (e.g. they mainly spent their time exploring the room and did not display behaviours indicative of distress or suggesting that they wanted to leave the room).

In the first experiment, each dog (n = 20 with 10 puppies aged 2-5 months and 10 adult dogs aged 13-48 months, from the Bidawee shelter; see Supplementary Table 2 for details) was tested during two successive playback sessions with: a) a ~30 seconds sequence of puppy-directed speech, and b) a ~30 seconds sequence of a human-directed speech (control). These two sequences came from our recording data bank (see Methods, part (i) Human speech recording and analysis) and were made of three successive renditions of the sentence: “Hi! Hello cutie! Who’s a good boy? Come here! Good boy! Yes! Come here sweetie pie! What a good boy!”. The playback sequences were recorded from the same human speaker for each dog, but each dog was tested with a different speaker. The two playback trials were separated by 1 to 2 minutes of silence, as the second playback was conducted once the dog had stopped displaying interest towards the speaker for at least 1 minute. Five puppies and 5 adult dogs heard the puppy-directed speech recording first while the other individuals heard the human-directed speech (control) signal first.

Because adult dogs from an animal shelter may have an unknown history of negative interactions with humans, we performed an additional set of trials on a sample of adult dogs kept as family pets and without history of re-homing (see Supplementary Table 2 for details). These dogs were tested using the same experimental setup as for the shelter dogs (design and size -3.5*4 m- of the experimental room, playback apparatus and protocol) and performed at the ENES Laboratory, Saint-Etienne (France), in September-October 2016. To ensure familiarity with the local language we use the following script: “Alors le chien! Comment ça va le doudou ? C’est qui le bon chien ? Viens ici mon chien ! Ah il est gentil le chien. Ca c’est un gentil chien !” recorded from 10 French native speaking female participants using the exact same protocol and material as with the US participants.

In the second experiment, each dog (n=10 puppies, aged 3-8 months, different individuals from those tested in the first experiment, see Supplementary Table 2 for details) was tested during two successive playback sessions with: a) a ~30 seconds sequence of puppy-directed speech, and b) a ~30 seconds sequence of adult dog-directed speech. These two sequences were derived from our recording data bank and were different for each tested dog. The two playback sessions were separated by 1 to 2 minutes of silence. 5 individuals heard the puppy-directed speech first while the other 5 individuals heard the adult dog-directed speech sequence first.

The experimental signals were played back through a Bose SoundLink Mini Bluetooth speaker II. This high-quality loudspeaker allows a faithful reproduction of human voice (see Supplementary Figure 1 for a comparison between the original and played back signals). The loudspeaker was positioned on the ground, near a corner and facing the centre of the room. The experimenter remained motionless, in the corner of the room opposite to where the loudspeaker was, and not facing the dog in order to avoid conscious or unconscious cueing. A video camera was placed to record the tested dog's reaction to the playback. The dog's response was assessed using the 11 following behavioural measurements (see Supplementary Methods). Instead of separately analyzing the dependent behavioural measures, we performed a principal component analysis (PCA) and retained a single composite score (PC1), separately for each of the two experiments [28] (Supplementary methods).

3. Results

(i) Human speakers use dog-directed speech with dogs of all ages

The analysis of recordings showed that dog-directed speech differs from control speech in both its spectral and temporal dimensions: 11 out of the 17 measured acoustic features were significantly affected by recording conditions (Supplementary Table 3). Specifically, dog-directed speech was higher-pitched, with more pitch variation over time. The periodic quality of the signal was also affected: harmonicity - the ratio of harmonics to noise in the signal - was higher in dog-directed speech sequences (Figure 1, Supplementary sound 1). Although human speakers modified their speech in front of dogs of all ages, post-hoc comparisons between recording conditions underlined that the distinctive pitch used in pet-directed speech was enhanced when speaking to puppies (Supplementary Table 3): in this condition our human speakers increased their mean pitch by 21% on average compared to normal speech (compared with 11% and 13% average increases when they spoke to adult and to old dogs respectively).

(ii) Only puppies are highly responsive to dog-directed speech

Results of the first series of playback experiments showed that speech quality, dog age, playback order as well as the interaction between speech quality and dog age were significant predictors of dogs' response to speech sequences (Table 1, Figure 2). As a result, 9 out of the 10 tested puppies responded more to puppy-directed speech than to human-directed speech, by reacting more quickly, looking more often at the loudspeaker and approaching it closer and for longer periods (Tukey post-hoc test on PC1 behavioural score: Z=3.34, P=0.0009, N=10; Supplementary Table 4 for loadings of behavioural variables on PC scores). Moreover, results of the second series of playback experiments showed that puppies did not respond significantly more to puppy-directed than to adult dog-directed speech (GLM: $\chi^2 = 0.44$ df = 1, $P = 0.509$), demonstrating that both types of dog-directed speech have similar stimulating effects.

In the first series of playback experiments, adult dogs responded less strongly to dog-directed speech sequences than puppies did (Tukey post-hoc test: Z=6.45, P<0.001, N=20 adult dogs and 10 puppies). Moreover, the behavioural response of adult dogs did not differ significantly between the two speech types, with 11 out of 20 individuals responding more to the dog-directed speech and the 9 others responding more to the human-directed speech (Tukey post-hoc test on PC1 behavioural score: Z=-0.37, P=0.708, N=20). The origin (shelter or family) of the tested dogs did not influence their behavioural responses ($\chi^2 = 0.45$, df = 1, P = 0.500, GLM with dependent variable = adult dog's behavioural reaction, fixed factors = speech quality, playback order and dog origin, random effect = dog identity).

(iii) Speech pitch is an important factor driving puppy behavioural response

As shown by the above acoustic analyses, human- *versus* dog-directed speech types differed with regards to several acoustic features. Assessing the impact of each of these features on dogs' behavioural reaction to playback reveals that there is a strong interaction between the effect of the mean pitch of the speech sequence and the effect of dog age (analysis restricted to dogs tested with English-spoken sentences: LME on PC1 scores of the first series of playback experiments, with playback order and interaction between pitch and dog's age as fixed effects and dog identity as random effect: $\chi^2 = 10.4$, df = 1, $P = 0.0012$; Figure 3; see also Supplementary Table 5 for interaction effects between other acoustic features and dog's age). Puppies' reactions were strongly influenced by the average pitch of the playback speech sequence: there was a highly significant effect of this acoustic feature on the level of behavioural reaction (LME on PC1 score of puppies with mean pitch and playback order as fixed effects and dog identity as a random factor: $\chi^2 = 11.0$, df = 1, $P < 0.001$, Figure 3). Conversely, the behavioural reaction of adult dogs to the playback was not significantly influenced by the pitch of speech sequence ($\chi^2 = 0.64$, df = 1, $P = 0.422$, Figure 3).

Two additional acoustic features significantly correlated with puppies' reaction to playback, albeit to a lesser extent than pitch: the percentage of the signal that is characterized by a detectable pitch (%voiced) and the harmonicity (harm) (Supplementary Table 6, Supplementary Figure 2).

4. Discussion

By showing that human speakers employ dog-directed speech to communicate with dogs of all ages, the present study suggests that this particular register of speech is used to engage interaction with a non-speaking, rather than just a juvenile listener. Yet dog-directed speech appeared to be modulated as expected by the "baby schema" hypothesis [14,15], as specific acoustic traits were further exaggerated when speaking to a puppy. At the receiver end, our playback experiments constitute the first demonstration that dog-directed speech functions to engage the attention of puppies, which are specifically sensitive to acoustic parameters as a higher mean pitch and a higher level of harmonicity. This speech pattern thus constitutes a functional signal promoting human-puppy interaction. Conversely, adult dogs displayed no significantly different preference for dog-directed speech, suggesting that this register loses its functional value in adult dogs.

The analysis of the acoustic structure of recorded sentences underlines differences between dog-directed and normal speech. In line with previous studies [26], we found that dog-directed speech is characterised by a higher pitch and a higher degree of harmonicity than normal speech. The fact that the visual presentation of dogs of all ages led human speakers to modify their speech pattern, is consistent with the hypothesis that dog-directed speech functions to facilitate interacting with an animal expected to be more sensitive to the prosodic, rather than to the verbal content of speech. Whereas caregivers progressively stop using infant-directed speech when infants start demonstrating syntactic and words understanding as they acquire language ability [31], human speakers continue using dog-directed speech with adult dogs that do not acquire language abilities. Pet-directed speech is thus in accordance with the "hyperspeech" hypothesis which states that speakers use speech patterns optimized for intelligibility [30]. In the case of dogs, this strategy may be efficient to promote word learning, an ability well-demonstrated in dogs [32].

The comparison of the acoustic structure between puppy-directed, adult dog-directed and old dog-directed speech recordings reveals that the age of the dog does weakly modulates the speech pattern: human speakers further raised the pitch of their voice when speaking to puppies than when speaking to adult and old dogs. The morphological cues typical of puppies (the “baby schema”) may thus constitute a reinforcing releaser. This effect of the “baby schema” could be further tested by assessing if people also change their speech pattern depending on the neotenic level of adult dogs, which varies among breeds [33].

As shown by playback experiments, puppies reacted strongly to dog-directed speech, demonstrating the functional value of this speech pattern. Whether this inter-specific dimension is innate or acquired through learning remains an open question. It is indeed well established that acoustic signals coding for emotional states share similar acoustic features across mammalian species [34]: although inter-specific communication may suffer from limitations [35-37], emotion-dependent similarities may derive from shared, ancestral production constraints or reflect convergent evolution in response to common selection pressures [38]. Dogs and wolves emit high-pitched tonal vocalizations in greeting contexts, between adults or between cubs, and as a solicitation for food or care [39], and it is likely that puppies are innately receptive to any high-pitched signals with a pronounced harmonicity. It is also likely that this innateness preference for pet-directed speech has been promoted by artificial selection: when choosing their pet within a litter, people will usually prefer puppies demonstrating higher levels of responsiveness to human solicitation [40]. Yet, this innate receptivity may also be reinforced by learning. The puppies we tested in our experiments had significant experience with humans and were used to interact positively with people who used dog-directed speech. It is indeed well established that dogs have a well-developed ability to associate prosodic cues of human speech with specific contexts [41, 42].

The absence of preferential reactivity to dog-directed speech in adult dogs was rather unexpected, as our production experiments suggest that old dogs are also exposed to humans using this speech pattern. This observation could be linked to an overall reduced propensity in adult dogs to respond to human playful signals. Specifically, in the absence of other communication cue (e.g. gestural signals), adult dogs could habituate rapidly to speech utterances from unknown persons, and thus rapidly ignore their vocal solicitation. Adult dogs are indeed known to react preferentially to their owner rather than to unfamiliar persons, although this depends on the context [43]. While puppies may react to any unknown speaker using pet-directed speech, older dogs may need additional cues to respond in unfamiliar contexts. Alternatively, this observation may suggest that pet-directed speech exploits perceptual biases which are present in puppies but not in adult dogs.

A potential limitation of our study arises from the fact that, in order to standardise the content of the dog directed speech utterances (see methods), we asked participants to read a script in front of pictures, which may have limited the extent of some features specific of dog-directed speech. Any such effect would however have been limited as we report clear differences between dog- and human-directed speech, both at the level of the acoustic properties, and at the level of the behavioural reaction that these utterances trigger in dogs. To address this potential limitation, future investigations could use stimuli recorded in a more realistic and interactive set-up, with participants asked to speak to “real” dogs instead of pictures.

In conclusion, while pet-directed speech appears to have some functional value in the context of human-puppy interaction, human speakers also use this speech format when speaking to older dogs, in spite of the absence of specific reactivity. This observation is consistent with the hypothesis that pet-directed speech is also a spontaneous attempt to get the attention of non-verbal, rather than just juvenile listeners. Dogs share many aspects of their “social competence” with humans [44] which causes dogs to appear ‘infant-like’ or ‘human-like’. The present study suggests that dogs may appear as mostly non-verbal companions to humans who consequently modify their speech features as they use to do when speaking to young infants. Such a speaking strategy seems to be employed in other contexts where the speaker feels, consciously or unconsciously, that the listener may not fully master language or has difficulty in speech intelligibility, such as during interactions with elderly people [45], or when speaking to a linguistic foreigner [47].

Ethics. All procedures described in this manuscript were conducted in accordance with appropriate USA and French national guidelines, permits and regulations, and the guidelines for the treatment of animals in behavioural research and teaching of the Association for the Study of Animal Behaviour (ASAB). Ethical approval for the playback experiments on dogs was granted by the Hunter College IACUC committee which waived the review requirement for this project. Experiments in France were done under Approval n°C42-218-0901 (ENES lab agreement, Direction Départementale de la Protection des Populations, Préfecture du Rhône).

Data accessibility. Additional methods, two figures, six tables and one sound recording are included as the electronic supplementary material.

Authors’ contribution. TBA, DR and NM designed the study; TBA and MGA performed the recordings and the playback experiments. All authors agree to be held accountable for the work performed.

Competing interest. The authors have no competing interests.

Funding. This study was supported by the Hunter College, City University of New York, and the University of Lyon/Saint-Etienne.

Acknowledgements. We would like to thank all participants, the Bideawee animal shelter, the owners of the dogs and Nicolas Boyer. NM was invited professor at Hunter College during part of this research project and is grateful to Chris Braun and Mark Hauber for their kind support.

References

1. Grieser DL, Kuhl PK. 1988 Maternal speech to infants in a tonal language: support for universal prosodic features in motherese. *Dev. Psychol.* **24**, 14–20.
2. Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskina VL, Stolyarova EI, Sundberg U, Lacerda F. 1997 Cross-language analysis of phonetic units in language addressed to infants. *Science* **277**, 684-686.
3. Cristia A, Seidl A. 2011 The hyperarticulation hypothesis of infant-directed speech. *J. Child Lang.* **41**, 913-934.
4. Cooper RP, Abraham J, Berman S, Staska M. 1997 The development of infants' preference for motherese. *Inf. Behav. Dev.* **20**, 477-488.
5. Naoi N, Minagawa-Kawai Y, Kobayashi A, Takeuchi K, Nakamura K, Yamamoto JI, Kojima S. 2012 Cerebral response to infant-directed speech and the effect of talker familiarity. *NeuroImage* **59**, 1735-1744.
6. Ko ES. 2012 Nonlinear development of speaking rate in child-directed speech. *Lingua* **122**, 841-857.
7. Werker JF, Pons F, Dietrich C, Kajikawa S, Fais L, Amano S. 2007 Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* **103**, 147-162.
8. McMurray B, Kovack-Lesh KA, Goodwin D, McEchron W. 2013 Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition* **129**, 362-378.
9. Ma W, Golinkoff RM, Houston D, Hirsh- Pasek K. 2011 Word learning in infant- and adult-directed speech. *Lang. Lear. Dev.* **7**, 209-225.
10. Liu, HM, Tsao FM, Kuhl P. 2009 Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and fiveyear-old children: a

longitudinal study. *J. Child Lang.* **36**, 909–922.

11. Golinkoff RM, Can DD, Soderstrom M, Hirsh-Pasek K. 2015 (Baby)talk to me: The social context of infant-directed speech and its effect on early language acquisition. *Curr. Dir. Psychol. Sci.* **24**, 339-344.
12. Narayan CR, McDermott LC. 2016 Speech rate and pitch characteristics of infant-directed speech: Longitudinal and cross-linguistic observations. *J. Acoust. Soc. Am.* **139**, 1272-1281.
13. Konner, M. 2010 The Evolution of Childhood: Relationships, Emotion, Mind (Belknap Press).
14. Lorenz K. 1943 Die angeborenen Formen moeglicher Erfahrung. *Z. Tierpsychol.* **5**, 235–409.
15. Glocker ML, Langleben DD, Ruparel K, Loughead JW, Valdez JN, Griffin MD, Sachser N, Gur RC. 2009 Baby schema modulates the brain reward system in nulliparous women. *Proc. Natl. Acad. Sci. USA* **106**, 9115-9119.
16. Hare B, Brown M, Williamson C, Tomasello M. 2002 The domestication of social cognition in dogs. *Science* **298**, 1634-1636.
17. Hare B, Tomasello M. 2005 Human-like social skills in dogs? *Trends Cogn. Sci.* **9**, 439–444.
18. Topal J, Gergely G, Erdohegyi A, Csibra G, Miklosi A. 2009 Communication in dogs, wolves, and human infants. *Science* **325**, 1269-1272.
19. Udell MAR, Dorey NR, Wynne CDL. 2010 What did domestication do to dogs? A new account of dogs'sensitivity to human actions. *Biol. Rev.* **85**, 327–345.
20. Rossano F, Nitzschner M, Tomasello M. 2014 Domestic dogs and puppies can use human voice direction referentially. *Proc. R. Soc. B* **281**: 20133201.
21. Müller CA, Schmitt K, Barber ALA, Huber L. 2015 Dogs can discriminate emotional expressions of human faces. *Curr. Biol.* **25**, 601-605.
22. Gregory JD. 2010 Pet custody: Distorting language and the law. *Fam. Law Quat.* **44**, 35-64.
23. Stoeckel LE, Palley LS, Gollub RL, Niemi SM, Evins AE. 2014 Patterns of brain activation when mothers view their own child and dog: An fMRI study. *PLoS ONE* , **9**, e107205.
24. Nagasawa M, Murai K, Mogi K, Kikusui T. 2011 Dogs can discriminate human smiling faces from blank expressions. *Anim. Cogn.* **14**, 525-533.
25. Hirsh-Pasek K., Treiman R. 1982 Doggerel: Motherese in a new context. *J. Child Lang.*

9, 229-237.

26. Burnham D, Kitamura C, Vollmer-Conna U. 2002 What's new, pussycat? On talking to babies and animals. *Science* **296**, 1435.
27. Boersma P, Weenink, D. 2012 Praat: doing phonetics by computer (Version 5.3.23) [Computer program].
28. McGregor, PK. 1992 Playback and Studies of Animal Communication (ed PK McGregor). Springer.
29. Barr DJ, Levy R, Scheepers C, Tily HJ. 2013 Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255-278.
30. Martin A, Utsugi A, Mazuka R. 2014 The multidimensional nature of hyperspeech: Evidence from Japanese vowel devoicing. *Cognition* **132**, 216-228.
31. Gogate L, Maganti M, Bahrick LE. 2015. Cross-cultural evidence for multimodal motherese: Asian-Indian mothers' adaptive use of synchronous words and gestures. *J. Exp. Child Psychol.* **129**, 110-126.
32. Kaminski J, Call J, Fischer J. 2004 Word learning in a domestic dog: evidence for ''fast mapping''. *Science* **304**, 1682–1683.
33. Borgi M, Cirulli F. 2016 Pet face: Mechanisms underlying human-animal relationships. *Front. Psychol.* **7**, 298.
34. Briefer E. 2012 Vocal expression of emotions in mammals: Mechanisms of production and evidence. *J. Zool.* **288**, 1-20.
35. Lingle S, Riede T. 2014 Deer mothers are sensitive to infant distress vocalizations of diverse mammalian species. *Amer. Nat.* **184**, 510-522.
36. Scheumann M, Hasting AS, Kotz SA, Zimmermann E. 2014 The Voice of Emotion across Species: How Do Human Listeners Recognize Animals' Affective States? *PLoS ONE* **9**, e91192.
37. Kelly T, Reby D, Levréro F, Keenan S, Gustafsson E, Koutseff A, Mathevon N, *in prep*. Adult human perception of distress in the cries of bonobo, chimpanzee and human infants. *Biol. J. Linnean Soc*.
38. Morton ES. 1977 On the occurrence and significance of motivation-structural rules in some bird and mammals sounds. *Amer. Nat.* **111**, 855-869.
39. Taylor AM, Ratcliffe VF, McComb K, Reby D. 2014 Auditory communication in domestic dogs: Vocal signalling in the extended social environment of a companion animal. In *The Social Dog, Behavior and Cognition* (eds J Kaminski, S Marshall-Pescini), pp. 131-163. Elsevier.

40. Riedel J, Schumann K, Kaminski J, Call J, Tomasello M. 2008 The early ontogeny of human–dog communication. *Anim. Behav.* **75**, 1003–1014.
41. Scheider L, Grassmann S, Kaminski J, Tomasello M. 2011 Domestic dogs use contextual information and tone of voice when following a human pointing gesture. *PLoS ONE* **6**, e21676.
42. Ratcliffe VF, Reby D. 2014 Orienting asymmetries in dogs'responses to different communicatory components of human speech. *Curr. Biol.* **24**, 2908-2912.
43. Kerepesi A, Doka A, Miklosi A. 2015 Dogs and their human companions: The effect of familiarity on dog-human interactions. *Behav. Process.* **110**, 27-36.
44. Miklosi A, Topal J. 2013 What does it take to become 'best friends'?Evolutionary changes in canine social competence. *Trends Cogn. Sci.* **17**, 287-294.
45. Masataka N. 2002 Pitch modification when interacting with elders: Japanese women with and without experience with infants. J. Child Lang. **29**, 939-951.
46. Uther M, Knoll MA, Burnham D. 2007 Do you speak E-N-G-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Com.* **49**, 2-7.

Table 1. Effect of speech quality (human-directed versus puppy-directed), dogs' age and order of playback on dogs' behavioural reaction to speech sequences.

term	**estimate**	**s.e.**	X^2	***df***	***p***
speech quality	-1.198	1.517	15.68	4	**0.0035**
dog's age	-1.860	1.938	29.79	4	**<0.0001**
playback order	-2.357	0.711	18.96	4	**0.0008**
speech*age	4.189	2.606	12.22	2	**0.0022**
speech*order	0.700	0.967	0.62	2	0.733
order*age	1.621	1.226	2.50	2	0.287
speech*order*age	-0.976	1.660	0.38	1	0.536

Figures legends

Figure 1. Influence of recording condition on speech quality. X-axis = recording conditions (directed speech to human adult, puppy, adult and old dog respectively). Y-axis = mean pitch of the recorded speech sequence. Each dot represents a single recording of the same speech sequence from different human adult speakers (each speaker was recorded in each of the four recording conditions; see main text for description of the recorded speech sequence). The size of dots is proportional to the degree of acoustic periodicity (ratio of harmonics to noise in the signal) of the recorded speech sequence. Violin plots show the distribution's density and dots are jittered horizontally for better visualization.

Figure 2. Dogs' behavioural reaction to playback of speech sequences. X-axis = dogs' age in months (logarithmic scale); Y-axis = dogs' behavioural reaction (represented as a Principal Component score PC1 calculated from 11 different behaviours; higher values mean stronger reaction to the playback signal). Each dot represents the result of one playback test. Each dog has been tested with two different speech qualities (red squares: reaction to puppy-directed speech; blue dots: reaction to human-directed speech). Solid lines = loess regression curves (degree of smoothing = 1; degree of polynomial = 1); grey shaded areas = confidence intervals.

Figure 3. Influence of speech pitch on dogs' behavioural reaction to playback. X-axis = mean pitch of the played-back sequence; Y-axis = dogs' behavioural reaction represented as a Principal Component score PC1 (higher values mean stronger reaction to the playback signal). Green triangles: reactions of puppies (aged 2-5 months); brown lozenges: reactions of adult dogs (aged 13-48 months). Solid lines = linear fits; grey shaded areas = confidence intervals.

Supplementary materials

Supplementary Table 1. Characteristics of the dogs presented during the recording of human speakers.

Supplementary Table 2. Individual characteristics of the dogs tested during the playback experiments.

Supplementary Table 3. Effect of recording condition on speech acoustic features.

Supplementary Table 4. Factors loadings on the Principal Components behavioural score.

Supplementary Table 5. Interaction effects between acoustic features and age on dog's behavioural reaction to speech sequences (*p* values are not Bonferroni corrected).

Supplementary Table 6. Effects of acoustic features on puppies' behavioural reaction to speech sequences (*p* values are not Bonferroni corrected).

Supplementary Figure 1. Narrowband waveforms and spectrograms of the original vocal sequence *(a)* and of the corresponding playback signal *(b)*. The played back signal was re-recorded at one meter from the Bose SoundLink Mini Bluetooth speaker II, in the room where dogs were tested. The spectrograms of both sounds are remarkably similar, with no noticeable filtering by the loudspeaker.

Supplementary Figure 2. Influence of *(a)* the percentage of the signal that is characterized by a detectable pitch (%voiced) and *(b)* the speech harmonicity (harm) on puppies' behavioural reaction to playback. The puppies' reaction is represented as a Principal Component score PC1 (higher values mean stronger reaction to the playback signal). Solid lines = loess regression curves (degree of smoothing = 2; degree of polynomial = 1); grey shaded areas = confidence intervals.

Supplementary sound 1. Examples of speech sequences recorded from the same human speaker. Successively: Human-directed (control), puppy-directed, adult dog-directed and old dog-directed.