Animal Behaviour 91 (2014) 126–134

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Cross-modal discrimination of human gender by domestic dogs

V. F. Ratcliffe, K. McComb, D. Reby*

School of Psychology, University of Sussex, Falmer, U.K.

ARTICLE INFO

Article history:
Received 19 September 2013
Initial acceptance 31 October 2013
Final acceptance 14 February 2014
Published online
MS. number: 13-00785R

Keywords:
categorization
cross-modal association
domestic dog
gender
social experience
vocal communication

We spontaneously categorize people as male or female, and when hearing a human voice we expect to see an appropriate sex-matched visual image. The extent to which domesticated species, which share our social environment, spontaneously develop such categorization abilities remains underinvestigated. Here we used a cross-modal preferential looking design to determine whether domestic dogs, *Canis familiaris*, spontaneously attribute an unfamiliar voice to a person of the corresponding sex. Fifty-one dogs were played a prerecorded male or female voice in the presence of a man and a woman. The responses were scored as correct or incorrect from both the direction of the first look and the total gaze duration towards each person after the voice presentation. Dogs living with one adult, or one man and one woman, performed significantly below chance as more (71%) of these dogs looked towards the incorrect person first. However, dogs living with more than two adults (including at least one man and one woman) performed significantly better, and significantly more (80%) of these dogs looked at the correct person for longer than they looked at the incorrect person. This suggests that while all of the dogs had spontaneously learnt to categorize human gender across sensory modalities, this ability was expressed differently depending on their social experience with humans. Dogs with greater experience, through regular exposure to multiple male and female human exemplars, responded by orienting towards the correct person, while those with more limited experience avoided looking towards the correct person. We discuss the importance of experience in determining the way that individuals spontaneously form and express categorization abilities.

Categorization is a key cognitive mechanism that determines how we perceive and process sensory information. As well as simplifying processing requirements (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), organizing stimuli into categories allows general inferences to be made and applied to new category members. Humans readily form complex, hierarchical categories representing their environment, using language to create specific referents that can coordinate categories between individuals (see Steels & Belpaeme, 2005 for a review). Currently, only a small number of studies have explored spontaneous category formation in other species, focusing on nonhuman primates (e.g. Murai, Tomonaga, Kamegai, Terazawa, & Yamaguchi, 2004; Murai et al., 2005). Comparative investigations into spontaneous category formation in nonhuman animals are therefore necessary to determine the functional relevance of this cognitive process in a broader range of species.

Domestic dogs provide an interesting model species to compare natural category formation in animals and humans. Dogs have shared the same environment as humans for at least 15 000 years (Savolainen, Zhang, Luo, Lundeberg, & Leitner, 2002), during which time they are likely to have undergone selection promoting specific sociocognitive abilities that allow effective cooperation and communication between the two species (Bräuer, Kaminski, Riedel, Call, & Tomasello, 2006; Hare, Brown, Williamson, & Tomasello, 2002). Added to this evolutionary predisposition is the effect of experience, as many dogs are extensively socialized with people, often sharing the same living habitat from an early age. As the human environment has become functionally relevant to dogs, this species may be expected to form spontaneous categories that are directly comparable with human categories.

It has already been established that, with training, dogs show equivalent categorization abilities to other mammals and birds. They are able to discriminate between 'dog' and 'nondog' sounds (Heffner, 1975), images of dogs and landscapes (Range, Aust, Steurer, & Huber, 2008) and images of dogs and other species (Autier-Dérian, Deputte, Chalvet-Monfray, Coulon, & Mounier, 2013), correctly generalizing their responses to novel stimuli. Spontaneous, ecologically relevant category formation is also evident in the dog's ability to form cross-modal perceptual associations when responding to familiar people. Using an expectancy

* Correspondence: D. Reby, School of Psychology, University of Sussex, Falmer BN1 9QH, U.K.
E-mail address: d.reby@sussex.ac.uk (D. Reby).

http://dx.doi.org/10.1016/j.anbehav.2014.03.009
0003-3472/

violation paradigm, Adachi, Kuwahata, and Fujita (2007) presented dogs with a photograph of either their owner's or a stranger's face after playing back one of their voices. Dogs looked for longer when the face did not match the preceding voice than when the stimuli did match, suggesting that dogs can use cross-modal associative categories when responding to familiar humans. This form of categorical perception is likely to be expressed naturally by dogs as the need to identify familiar humans has a clear function in recognizing important social partners and care providers. The spontaneous use of an interspecific category representing familiar humans leads to the possibility that it may also be relevant for dogs to form categories about unfamiliar humans, which would allow direct comparisons with our own categories.

One of the predominant ways that we categorize unfamiliar people is by their gender, primarily by associating visual and vocal cues. Because human faces are sexually dimorphic (Burton, Bruce, & Dench, 1993), differing in both shape and texture (Hill, Bruce, & Akamatsu, 1995), face gender classification in adults is close to 100% accuracy (O'Toole et al., 1998). Sexual dimorphism also leads to differences in the vocal tract anatomy of adult men and women. The larger adult male larynx results in a difference of approximately 80 Hz in fundamental frequency (F0) between the voices of adult men and women, with mean values at around 120 Hz and 200 Hz, respectively (Titze, 2000). Additionally, adult men have a disproportionally longer vocal tract than women (Vorperian et al., 2009), causing lower first formant (F1) values and formant dispersions approximately 15–20% lower than in women (Fant, 1960; Goldstein, 1980). The relative F0 and formant values classify the gender of adult voices at 98.8% accuracy (Bachorowski & Owren, 1999). The presence of both visual and vocal gender cues enables cross-modal perceptual matching of voices to individuals from an early age (Walker-Andrews, Bahrick, Raglioni, & Diaz, 1991).

Dogs are also likely to be able to perceive these gender differences in the human voice, as they attend to variation in formants to determine size information in conspecific vocalizations (Taylor, Reby, & McComb, 2011) and can be trained to discriminate between average male and female F0 differences in human vowel sounds (Baru, 1975). Gender-specific behavioural differences in the way humans interact with dogs have been identified (e.g. Prato-Previde, Fallani, & Valsecchi, 2005) which could have created the need for dogs to categorize human gender in order to adjust their responses appropriately. In support of this, shelter-housed dogs petted by women show more relaxed behaviour and lower cortisol levels than those petted by men (Hennessy, Williams, Miller, Douglas, & Voith, 1998), and are more likely to direct defensive aggressive behaviour towards men than women (Lore & Eisenberg 1986; Wells & Hepper, 1999). Although the specific cues to which the dogs were responding cannot be determined from these studies, they do suggest that categorically assessing human gender could be functionally relevant for dogs, influencing their reaction to the individual person. Therefore the ability to perceive and associate different sensory cues to human gender as categorically equivalent would be a useful ability.

To determine whether dogs do categorize human gender using different sensory cues, we tested whether they associate voices with unfamiliar people using gender cues in a cross-modal preferential looking paradigm, in which subjects were required spontaneously to match voices to people by their gender. In our study, a man and a woman stood either side of a loudspeaker from which a voice recording of a different person was played. Dogs were positioned facing the centre line, and their visual orientation to the person matching the gender of the voice and the nonmatching person were recorded. If dogs spontaneously combine vocal and visual cues to identify human gender cross-modally, it was predicted that they would look first, and for a longer duration, at the person of the same gender as the voice. The potential effect of social factors on performance was also investigated, as well as possible mechanisms involved in such variation.

METHODS

Subjects

A total of 51 adult dogs of 17 different breeds were recruited when their owners responded to advertisements in the East Sussex area. Ages ranged from 7 months to 11 years old (mean + SD = 5.03 + 3.17 years), including 26 males and 25 females. The selection criteria for subject animals were that they had to be healthy adults (older than 6 months) with no known sight or hearing problems and no known aggression towards people. Subjects and their owners were naïve to the experimental set-up and had not participated in any previous vocal communication or behavioural research.

Playback Acquisition

Nine men and nine women, aged between 20 and 52 years (mean + SD = 30.94 + 9.75 years), were audio recorded after being instructed to pronounce the following phrases as if speaking to a dog in a positive voice: 'Hey!', 'Come on then', 'Good dog!', 'What's this?'. Each speaker pronounced each phrase once. All recordings were made using a Zoom H4N Handy Recorder in a soundproof booth. The sampling frequency was set at 44 100 Hz, with a 32-bit sampling rate, for each recording. The vocal parameters of the recordings were then checked for a bimodal distribution according to gender using PRAAT v.5.0.3 (http://www.fon.hum.uva.nl/praat/). The four phrases were analysed together as a single audio file. The mean, minimum and maximum F0 values were calculated using the PRAAT autocorrelation algorithm 'to Pitch (ac)' which estimates the F0 contour across the utterance. The mean F0 for the male voices was between 142.00 Hz and 193.48 Hz (mean + SD = 166.80 + 17.64 Hz), while the mean F0 for the female voices was between 251.13 Hz and 405.99 Hz (mean + SD = 323.26 + 61.22 Hz). The F0 ranges (maximum F0–minimum F0) for the male voices were 109.70–164.73 Hz (mean + SD = 154.16 + 48.32 Hz), while the female F0 ranges were 269.41–528.01 Hz (mean + SD = 350.20 + 86.35 Hz). The formant dispersion (ΔF) was calculated using the PRAAT Linear Predictive Coding 'Burg' algorithm, which estimates the centre frequencies of the first four formants across the utterance. These values were then used to calculate the average spacing between the formants. The male ΔFs were between 927.98 Hz and 1120.60 Hz (mean + SD = 1029.15 + 71.39 Hz), while the female ΔFs were between 1140.60 Hz and 1241.00 Hz (mean + SD = 1215.56 + 45.20 Hz). All of the recordings were normalized to −1.0 dB maximum amplitude in Audacity 2.0.0 (http://audacity.sourceforge.net).

Experimental Set-up

Experiments were carried out between June and September 2012 at two indoor test locations in the East Sussex, U.K., area (The Dog Hut in Barcombe and Hamsey Riding School in Lewes). A cross-modal preferential looking paradigm was used. The design was developed on the basis of pilot trials conducted in April and May 2012 on 20 subjects, who did not take part in the final study trials. The original piloted study included a sequence of six trials per subject; however, we found that habituation to the procedure led to a reduction in responses after the first trial. Therefore in the full study each dog took part in only one trial.

An Anchor LIB-6000H Liberty loudspeaker (frequency response: 60 Hz–15 kHz) was mounted onto a 130 cm tall stand and