Understanding social communication between humans and dogs

Alex Victoria Benjamin

PhD

University of York
Psychology

November 2018

Abstract

This research investigates social communication between humans and dogs. Chapter 1 outlines the current understanding of dog social communication, highlighting outstanding questions with regards to the production and perception of speech to dogs, and the function of barking in domesticated dogs. The first study investigated individual differences in the production of a special speech type known as dog-directed speech (DDS), which is higher in pitch, pitch modulation, and affect compared with adult-directed speech (ADS). Results showed that females increase their pitch more than males in DDS, but that the difference in pitch modulation and content of speech between ADS and DDS does not differ between males and females. Neither experience with dogs, nor levels of empathy influenced production of DDS. The second study used playback experiments to investigate whether dogs prefer DDS compared to ADS and found that dogs display an attentive and social preference for DDS, but only when both content, and prosody of speech are dog-directed. This demonstrated for the first time that naturalistic DDS may function to improve the affiliative bond between humans and dogs. The third study found that direct eye-gaze selectively enhances the preference for DDS, suggesting that both speech type, and eye-gaze are important cues for communication with dogs. Finally, I examined whether dogs understand contextual information conveyed by conspecific barks. Findings suggest that dogs do not use contextual information in conspecific barks to inform their behaviour in a naturalistic setting and that humans are more sensitive to these cues than dogs. This suggests that increased barking in domestic dogs may have evolved as a means of communicating with humans, rather than conspecifics. Overall, this research demonstrates the complexity of dog-human communication and highlights the value of comparative research for gaining insight into the evolution of sophisticated social communication resulting from both natural and human selection.

List of Tables

List of Figures

Table of Contents

Acknowledgements

I have many people to thank for providing support and guidance throughout the past several years. Most importantly, thanks must go to my supervisor, Professor Katie Slocombe, who has supported me at every stage of my PhD, even at a time when my belief in myself had faltered. Under Katie's mentorship I have developed considerably as a researcher and learnt a great deal about social communication in animals, but Katie's support has been above and beyond anything I could have hoped for in a supervisor and I am extremely grateful for all her advice, personal support, and ever-positive attitude that have made these years of study so memorable.

I would also like to thank Alyse and the staff and dogs at the Redhouse Boarding Kennels, and Vanessa and her team, and all the dogs at the Waggy Dog Creche, for letting me conduct my research with them. Also thank you to the residents of Newton-upon-Derwent, York whose dogs provided valuable pilot data at various stages, and participated in my final bark study. I am also grateful to the research assistants who helped me, particularly Anna Auster for her expert dog handling, and Sophie Marshall for her reliability and enthusiasm even when it was cold and raining! Thank you to the undergraduate students Ella, Phoebe, Keira, Helen, and Tom for lending their voices to the stimuli, modelling and handling during data collection, and coding videos for reliability analyses.

Thanks to all of the fantastic dogs who participated, listened, wagged their tails, and generally made my PhD great fun. Particular thanks must go to Poppy, Lily, Zeus, Suzy, Teddy, Jake and of course, Cooper. Woof woof woof!

For their support and encouragement throughout my PhD, I thank my family and friends, particularly Aimee who always has a smile and a hug to share, and is the best for a coffee break and a gossip, and Freya; I am so glad we became so close and have shared so many fun experiences together including riding polo ponies and talking on BBC Radio York!

Finally, my love and gratitude to Sam, who has supported and encouraged me, been patient and kind, and made me laugh every day. I couldn't have done this without you.

Author's Declarations

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other University. All sources are acknowledged as References. Each of the studies was conducted in accordance with the ethical standards of the University of York departments of Psychology, and Biology.

Chapter 3: Dogs prefer naturalistic dog-directed speech has been published in the journal *Animal Cognition* as:

Benjamin, A., & Slocombe, K. (2018). 'Who's a good boy?!' Dogs prefer naturalistic dog-directed speech. *Animal Cognition*, 21(3), 353–364. https://doi.org/10.1007/s10071-018-1172-4

For this study I collected, coded and analysed all data. Katie Slocombe aided in the design, provided intellectual input and guidance throughout and contributed to manuscript edits.

Chapter 1: Introduction

The most widespread form of interspecies bonding is that which occurs between humans and dogs (Tacon & Pardoe, 2002). Dogs developed close contact with humans some 5000 years earlier than any other species (Kubinyi, Viranyi & Miklosi, 2007) and were first domesticated in Europe over 18,000 years ago (Thalman, Shapiro & Cui, 2013). In understanding the domestication process in dogs from wolves, two stages of selection have been proposed. The first stage occurs in the absence of intentional breeding by humans and is a form of natural selection often described as *self-domestication.* In this stage, less fearful and aggressive wolves had a selective advantage over their fearful and aggressive group mates, as they were able to approach human settlements to exploit opportunities such as scavenging for food and human waste (Coppinger & Coppinger, 2001). These 'proto-dogs,' with reduced fear and aggression and increased contact with humans, were able to interact with humans more pro-socially than wolves and may have returned to human camps several times to develop trusting relationships with the humans who then encouraged the dogs to join them. The second stage is direct, intentional breeding by humans for desirable characteristics. It is not thought that dogs were domesticated with the specific intention of any direct benefit for the humans (e.g. food, or hunting); instead, early dog fossils from human burial sites indicate that dogs shared a special, perhaps spiritual relationship with humans (Morey, 2006). Later selection of preferred appearance and particular characteristics displayed by these early dogs, such as hunting ability, speed, or sensitivity to scent, led to the emergence of the variety of dog breeds that we are familiar with today.

The self-domestication hypothesis

When investigating the process of domestication, it is useful to examine common behavioural traits. The main behavioural trait observed in domesticated species is a reduction in aggression, including intraspecific, interspecific, offensive and defensive aggression (Hare et al, 2012). The comparison of wolves and feral dogs illustrates this reduction clearly, and wolves and dogs are the best-studied pairing of wild ancestors and domesticated descendants. For example, for wolves, defence of territory is extremely important. Individuals try and avoid 'buffer-zones' where lethal and intense group encounters frequently occur between groups, with adult mortality due to intraspecific killing reaching 65% in some areas of the USA (Mech, 1994; Mech et al, 1998). In contrast, feral dogs rarely engage in physically aggressive disputes during interactions with conspecifics, despite similar group sizes compared with wolves (Boitani et al, 1995; Pal et al, 1999). Instead, feral dogs tend to bark at one another from a distance, intimidating the other individual until one of them chooses to leave (Bonanni et al, 2010; Macdonald & Carr, 2016; Pal et al, 1999). Similarly, intragroup aggression is also lower in dogs than in wolves, where female wolves are frequently observed engaging in infanticide as a result of reproductive competition, with dominant females killing the pups of subordinate females (McLeod, 1990, Mcleod et al, 1996). In feral dogs, there is little evidence of reproductive competitiveness, and no reports of infanticide among females (Pal et al, 1999). Furthermore, dogs also display increased tolerance compared with wolves, often accepting inspection of their anogenital areas by a stranger, where wolves are likely to become aggressive when approached in this way by another individual, even if the approach is by a group mate (Bradshaw & Nott, 1995). This down-regulation of aggression and increased social tolerance is thought to be a marker of self-domestication, in the absence of intentional human selection.

While the main behavioural trait consistent across domesticated species is reduced aggression, a suite of other behaviours or changes have also been observed to occur with domestication. These are often referred to as the traits of the domestication syndrome and can be seen in many different species (Wilkins, Wrangham & Fitch, 2014). Physiological changes such as lower reactivity of the hypothalamic-pituitary-adrenal (HPA) axis (Kunzl & Sachser, 1999), which is involved in the fight-or-flight response and linked to aggression, as well as increased frequency of reproductive cycles have been observed in domesticated species. Correlated changes in anatomy are also demonstrated, such as reduction in cranial capacity, shortening of the face, reduction in tooth size, floppy ears, and depigmentation of parts of the body (Hemmer 1990; Kruska, 2005; Hare et al, 2012), as well as behavioural changes including increased pro-social behaviours such as play, grooming, and non-conceptive sexual behaviours, which seem to be linked to social factors such as affiliative relations, relationship control/conflict resolution and mate selection (Furuichi, Connor & Hashimoto, 2014). It is not thought that these traits are the result of direct selection, rather these traits seem to have emerged as correlated by-products of selection against aggression.

It is perhaps reasonable to argue that behavioural traits of the domestication syndrome could have emerged as a result of selective breeding in captivity. However, Kunzl et al (2003) demonstrated that keeping wild species in captivity does not by itself reduce aggression. Domesticated guinea pigs raised in captivity show an attenuation of aggression, often allowing multiple males to be housed together, whereas offspring of wild species caught and bred in captivity showed no reduction in aggression or increase in socio-positive behaviours, suggesting that captive breeding alone is not sufficient for domestication to occur, and that selection for reduction in aggression by humans may be necessary. The longest running experimental investigation of this is in silver foxes in Siberia (Belyaev, 1979). An experimental population of silver foxes who had little interaction with humans were selectively bred for displaying low-reactivity towards humans at 7 months old (as demonstrated by the absence of bite attempts, and increase in approach behaviours). A control population was bred randomly with regards to their behaviour towards humans. Subsequent generations of experimental foxes displayed increased approach behaviours with humans instead of trying to bite or avoid humans when touched compared to controls. After 20 generations, most experimental foxes were friendly towards humans in ways that compare to domesticated dog puppies (including wagging their tails when approached by a human). The experimental foxes displayed evidence of domestication syndrome such as changes in HPA axis activation, flatter faces, floppy ears, narrower skulls, and submissive posture with high pitched vocalizations used in the presence of humans. Female foxes in the experimental group also became fertile biannually (like dogs) compared to annually (like wolves). As a result, fox kits from the experimental population were found to be as skilled as dog puppies at spontaneously using human communicative gestures, and were more skilled than control foxes at problem solving, even though this ability was not selectively bred for by humans. Moreover, genetic convergence between the experimental foxes and domesticated dogs were observed (Kukekova et al, 2006; 2011). These findings demonstrate that traits of the domestication syndrome emerge as correlated by-products of selection against aggression, and it is suggested that the mechanism by which the domestication syndrome emerges is though changes in development. For example, wolf puppies are able to be socialised with humans in the absence of conspecifics (Miklosi et al, 2003), whereas adult wolves are not. Dogs retain this window for socialisation far longer than wolves, and are therefore able to be socialized to a greater extent into adulthood. Wolves also display a wider array of antagonistic signals in adulthood than in puppyhood, while even adult dogs only display a small number of juvenile signals of this kind. The retention of juvenile characteristics in adult domestic animals suggests that ontogenetic changes may be responsible for the behavioural differences observed in wild and domestic species, and that these changes may be derived from delayed development of physiological responsiveness.

The domestication syndrome has been achieved through experimental selection for reduced aggression; however, whether this can be achieved in the absence of human interaction is not addressed in the study of animals in captivity. The evolution of dogs from wolves provides some insight into the domestication process in the wild, and provides a useful model for the study of other species pairs more closely related to humans. The self-domestication hypothesis developed from the study of wolves and dogs has been applied more recently to congeneric apes (Hare et al, 2012). Chimpanzees and Bonobos are closely related species that share phenotypic characteristics, but differ in their levels of aggression. Chimpanzees display severe intergroup aggression, as well as within their social groups, often using aggressive displays to compete for social status, resources, and mates, as well as to intimidate females and other group members (Muller et al, 2007). Aggression is more prevalent in males, with adolescent males asserting dominance over adult females, but it also occurs between females, particularly in relation to reproductive competition and infanticide (Pusey et al, 2008). Intergroup interactions are most often aggressive, and often include fatalities, particularly infants of outgroup females, and solitary adult males (Wrangham & Glowacki, 2012). In comparison, aggression in bonobos is far reduced, both within and between groups. While dominance is still displayed, bonobos display aggression with lower intensity, and rarely involve physical aggression with conspecifics (Furuichi, 1997; Furuichi et al, 1998). Although bonobos are also territorial, they are far less aggressive in intergroup interactions than chimpanzees, with only a small number of interactions resulting in physical violence (Hohmann & Fruth, 2002), instead one group will usually leave, as is the case with dogs. While displaying lower levels of aggression, bonobos are also more prosocial than chimpanzees, engaging more readily in social behaviours with conspecifics including play, grooming, and copulation (Furuichi, 2011). Along with reduction in aggression, bonobos also display phenotypic and developmental traits consistent with the domestication syndrome. For example, bonobos compared to chimpanzees display reduced cranium and tooth size, depigmentation of body parts including lips, and reduced sexual dimorphism. Bonobos also display evidence of peadomorphism, retaining juvenile traits such as white tail tufts, and paedomorphic cranium (Lieberman et al, 2007). Together, these traits of the domestication syndrome suggest that the self-domestication hypothesis developed in the study of wolf-dog domestication, is useful for studying the emergence of species in the wild without direct selection by humans. Hare et al (2012) suggest that a possible alternate theory explaining the differences between chimpanzees and bonobos, could be selection for severe aggression in chimpanzees from a bonobo-like ancestor, or that both chimpanzees and bonobos were derived from a common ancestor that possessed traits seen in both species. However, study of the skulls of both chimpanzees and bonobos suggests that the skulls of chimpanzees closely follow the ontogenetic patterns of skulls belonging to the more distant ancestor, the gorilla, compared to the bonobo skull which remains more juvenile. This suggests that chimpanzee and gorilla cranial ontogeny is ancestral, while the bonobo cranial ontogeny is derived (Lieberman et al, 2007), and that behavioural ontogeny in bonobos is also derived. This supports the explanation that reduction in aggression in bonobos is linked to self-domestication, where natural selection is the only possible mechanism for the down-regulation of aggression (Hare et al, 2012).

Social communication

Humans are able to build strong social relationships with dogs, in a way that is not possible with other non-domesticated animals. It is thought that this may be due in part to the superior communicative ability that dogs display in comparison to other species, particularly the closely related, non-domesticated wolf. Wolves that are highly socialized with humans from birth are able to follow touching cues and some pointing cues from familiar human experimenters in order to locate hidden food (Miklosi et al, 2003). Although more successful than non-socialized wolves, even wolves socialized intensely from birth perform significantly worse in these tasks than domesticated dogs. Further, during a simple manipulation task, when faced with an impossible trial, dogs voluntarily look towards a human for guidance or help, where even socialized wolves do not. Rather, the wolves give up when they cannot succeed by themselves, and resume other behaviours such as sleeping (Miklosi et al, 2003). Dogs, but not wolves are also able to engage in joint attention and cooperation with humans to achieve goal-directed actions (Range & Viranyi, 2014). Based on these findings, it is reasonable to argue that one important difference between dog and wolf behaviour lies in their ability and willingness to look towards humans for information – particularly the face.

In human communication systems, looking behaviour is important for initiating and maintaining communicative interactions. It is possible that the readiness of dogs to look at human faces may have led to the development of complex human-dog communication that cannot be achieved in wolves, even after extended and intense socialization. (Hare, Brown & Tomasello, 2002; Miklosi et al, 2003). Human infants display a preference for looking towards faces from a very early age, with evidence showing that newborn infants are able to follow a slowly moving face stimulus further than they follow scrambled face stimuli (Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis & Morton, 1991). This preferential tracking, along with evidence that infants also imitate facial expressions at a very early age, suggests that face perception is extremely important in the development of social communication in humans (Haxby, Hoffman & Gobbini, 2000). Initiation of communicative intention through eye-gaze cues is the basis for natural pedagogy (Csibra & Gergely, 2009) as it establishes communicative context (Kleinke, 1986; Hains & Muir, 1996; Symons, Hains & Muir, 1998) and provides a consistent cue to infants that a communicative signal is meant for them. This allows infants to attend to behaviours that are important for learning and development.

While the ability to use eye-gaze as an ostensive cue has been clearly demonstrated in human infants, the evidence for using gaze cues between humans and other species is less well established. Ferrari, Kohler, Fogassi and Gallese (2000) used physiological studies to show that neural correlates of eye direction detection exist in adult pig-tailed macaques who could follow the gaze of an experimenter using eye cues alone. This skill seems to develop over a lifetime, with juvenile monkeys unable to orient their attention on the basis of eye cues, but rather relying on a combination of eye and head cues to accurately follow gaze. Further research with capuchin monkeys and rhesus monkeys (Anderson et al., 1995; 1996) found that these monkeys could learn to discriminate objects based on human gestures such as pointing, but when relying on eye-gaze alone, the monkeys performed at chance level. This suggests that while non-human primates appear to use human eye-gaze cues in communication in some situations, they seem less sensitive to these cues in comparison with humans. This is likely partly due to direct gaze being perceived as a threat in most non-human primates, with dark schlera also making gaze more difficult to detect in conspecific interactions. Although direct gaze may be avoided in most non-human primates, the process of domestication may have resulted in domestic animals finding human gaze less threatening and humans may have selected for sensitivity to human communicative cues, including gaze, through the domestication process.

Not only do human infants display a preference for looking towards human faces, and process gaze-cues more reliably than non-human primates, there is also evidence of physiological changes during the sharing of mutual gaze in both adults and human infants. In mother-infant communication, it is possible to observe an increase in oxytocin when the duration of mutual eye-gaze is increased (Kim, Fonagy, Koos, Dorsett & Strathearn, 2014). Oxytocin plays a fundamental role in the mediation of affiliative bonding between mother and infant, and between partners of monogamous species (Young & Wang, 2004). Reciprocal communication is thought to involve an oxytocin positive loop, in which maternal oxytocin leads to increased maternal nurturing, which in turn leads to increased oxytocin in the infant (Nagasawa, Okabe, Mogi & Kikusui, 2012). Interestingly, it has been shown that dogs are able to hijack this oxytocin positive loop in cross-species relationships with humans. Nagasawa et al (2015) found that mutual eye-gaze with dogs increased oxytocin concentrations in humans, but that this was not the case when sharing gaze with wolves. Subsequently, humans displayed more affiliative behaviours towards the dogs, just like they did with human infants, which in turn increased dogs' oxytocin levels. They also found that administering oxytocin to dogs nasally, increased gazing behaviour in dogs, which also led to increased oxytocin levels in owners, suggesting a causal link between oxytocin and mutual eye-gaze. The authors argue that the cross-species oxytocin loop modulated by sharing mutual gaze explains the evolution of the human-dog relationship, as dogs are able to engage in common methods of communicating social attachment. This suggests that the domestication of dogs may have involved co-opting social-cognitive systems in humans that are fundamental for social attachment, usually found only in intra-species mother-infant relationships.

As discussed in relation to the self-domestication hypothesis, it is difficult to know whether domestic dogs' ability to read human gestures is the result of human selection or as a by-product of selection against fear and aggression towards humans (Hare et al, 2005; Belyaev, 1979). In order to understand which of these competing explanations is most likely, Hare et al (2005) used a similar paradigm as that of Belyaev (1979), to examine the socio-communicative abilities in an experimental population of fox kits selectively bred over 45 years to approach humans fearlessly and non-aggressively. The experimental foxes were able to reliably identify the location of hidden food following pointing, and gazing cues delivered by a human experimenter. These foxes performed as well as dog puppies on this task, with both foxes and puppies performing significantly above chance. Moreover the experimentally bred foxes performed significantly better compared with fox kits from a control population that were not bred for tameness, who performed at chance level. These results suggest that socio-communicative evolution occurred in the experimental foxes unintentionally as a result of selection for increased tameness and reduced aggression. It is likely therefore that higher sensitivity to human social cues evolved in dogs as a result of selection for tameness and non-aggression towards humans, rather than direct selection for improved socio-cognitive ability during the domestication process. Thus far, dogs' sensitivity to human visual communicative cues (gaze, manual pointing) have been discussed, but humans also communicate with each other and dogs with auditory signals.

Infant-directed speech

Parallels have been drawn between the way we talk to human infants and the way we address our pets, particularly dogs. When talking to an infant, adults use a special speech type characterised by elevated fundamental frequency (pitch), exaggerated intonation contours, and high affect (Burnham, Kitamura & Vollmer-Conna, 2002). It is thought that this speech type, known as infant-directed speech (IDS) facilitates infants' linguistic development, allows infants to select appropriate social partners (Schachner et al, 2011), and increases social bonding between the infant and caregiver (Kaplan et al, 1996).

Although it is thought that IDS functions to aid language acquisition, the mechanism by which this occurs is not yet clear. Thiessen, Hill, and Saffran (2005) believed that IDS facilitates word segmentation from fluent speech. It is known that infants are likely able to extract word boundaries using categorical perception (Eimas, 1971) based on the statistically systematic relationship between neighbouring speech sounds (Saffran, Aslin, and Newport, 1996). Thiessen et al, (2005) used a string of nonsense words spoken in IDS, and compared infant word learning between this and the same string of nonsense words spoken in adult-directed speech (ADS). Word learning was measured by the infant's ability to distinguish between speech sounds (phonemes) that occurred across a word-boundary using a habituation technique. Infants between 6 and 8 months old were able to distinguish between words across statistical boundaries after hearing IDS, but not after exposure to ADS. This suggests that the acoustic properties of IDS may be important for word segmentation in the acquisition of language.

It is thought that the linguistic benefit of IDS may also come from the amplification of the phonetic characteristics of native language vowels (Kuhl et al, 1997). It seems that when speaking to their infants, mothers do not simply raise the frequency of all speech formants, as would be the case if they were simply mimicking childlike speech. Instead, the difference between first, second, and third formants of IDS tends to be exaggerated. This is reflected acoustically in exaggerated /i/, /u/, and /a/ vowel sounds, and was originally thought to improve the clarity of speech. This hyperarticulation of vowel sounds does not only occur in IDS, but also when speaking to a foreign adult or second-language learner, who may have similar linguistic needs to an infant, but does not possess the same affective needs (Uther, Knoll, & Burnham, 2007). Uther et al, found that hyperarticulation is similar for IDS and foreigner-directed speech, but affect is lower for foreigner-directed speech than even ADS. This suggests that acoustic modifications to vowels are pedagogical and occur independently of changes in pitch and affect, thus supporting the idea that it is the hyperarticulation of IDS that facilitates its linguistic benefit. However recent research by Martin et al, (2015) has found that the phonetic distribution of the vowel sounds /i/, /u/, and /a/, overlap more in IDS than in ADS, leading to poorer performance of speech recognition algorithms. This suggests that if hyperarticulation functions to improve language learning, the mechanism underlying this benefit may be more complex than simply improving clarity.

It is thought that IDS is a robust cross-cultural phenomenon, with evidence for IDS across several languages including English, Japanese, Swedish, French, Russian, and German (Fernald et al., 1989). For instance, Fisher and Tokura (1996) studied IDS produced by English and Japanese mothers, identifying exaggerated intonation contours, vowel lengthening, and local acoustic changes in duration and pitch across the two languages. It is clear however, that the vast majority of studies have occurred in WEIRD (western, industrialized, educated, rich, and democratic) populations. For instance, there is very few studies of the properties of IDS in

African countries (Broesch & Bryant, 2015). Moreover, the majority of studies that discuss the acoustic properties of IDS do so with female speech. It is possible that this is due to the recruitment of mothers with infants, but there is far less evidence investigating whether men spontaneously produce IDS, and whether the acoustic properties differ between genders. One study by Fernald et al., (1989) acoustically analysed the speech of mothers and fathers in French, Italian, German, Japanese, British English, and American English. They found that mothers and fathers increased the pitch of their speech for infants compared with ADS, but highlight that mothers, not fathers, used a wider range of frequencies in IDS compared to ADS. They also found that pitch modulation was greatest in American English compared to other languages. This study suggests that while IDS may be produced in several languages, the acoustic properties of IDS in each language may not be as universal as previously suggested. The difference in speech by males and females also highlights the need for more extensive study of individual variation in production of IDS.

Dog-directed speech

In the same way that infant-directed speech (IDS) is produced automatically when talking to infants, in Western cultures humans also frequently produce a special speech type when talking to their pets. This pet-directed speech (PDS) shares some of the acoustic features of IDS including high frequency and exaggerated affect (Burnham et al, 2002). According to an experiment by Burnham et al (1998), pitch is significantly higher for IDS and PDS than ADS, but does not differ between IDS and PDS. Affect, measured by ratings of low-pass-filtered speech, in which intonation and rhythm can be heard, but words cannot be understood, was found to be highest in IDS, while the affect of PDS was significantly higher than that of ADS. Burnham et al., (2002) also found that participants hyperarticulated their vowel sounds when producing IDS, but not when producing PDS and ADS. From these results it is reasonable to suggest that pitch may be elevated in IDS and PDS in order to attract the listener’s attention, while affect is elevated to meet listener's emotional needs, possibly motivating affiliative interaction with the speaker. The fact that hyperarticulation of vowels occurs in IDS but not PDS, supports the idea that this may be the aspect of IDS that assists spoken language acquisition (Kuhl et al, 1997) and the speaker's hyperarticulation may be mediated by the perceived linguistic capacity of the receiver; evidence that supports this view is provided by a study that compared speech produced to dogs, parrots and infants. Speakers seem to hyperarticulate their vowels most with prelinguistic human infants, followed by parrots, with little evidence of this when addressing dogs, who in contrast to parrots have no ability to produce speech (Xu, Burnham, Kitamura & Vollmer-Conna, 2015). As dogs are unable to learn human language, hyperarticulation of vowels is unnecessary when addressing them, which may explain why naturally occurring dog-directed speech (DDS) does not contain this modification. This suggests that we unconsciously alter the acoustic properties of speech depending on the linguistic potential as well as the affective needs of a listener.

The study by Burnham et al (1998) is one of the earliest experiments exploring the differences between IDS and PDS and is highly cited in the literature as a result. However, the study analysed the speech of 12 participants when talking to their 6- month old infant, their pet, and another adult. In this way, the study is limited, as conclusions can only be drawn about interactions with a familiar listener. Caution should be employed when using this data to draw conclusions about ostensive communication, as the speech type an adult uses to talk to their pet/infant, may differ from the speech type an adult uses when talking to an unfamiliar animal/infant. Furthermore, it is important to note that the adult participants in this study were all females with infants. As mentioned above, this bias is true of many studies of IDS as well as the few studies that investigated PDS. It is reasonable to suggest that PDS produced by female adults may differ from that produced by other groups such as males, and children, and also that this data may not generalise to females who are not mothers. It is possible that IDS is modulated by hormonal changes that help strengthen the bond between a mother and her infant (Kaplan, Danko, Kalinka & Cejka, 2012), and PDS produced during this time may also differ to that produced by non-mothers. To address this Hirsh-Pasek et al. (1982) studied the speech of four women, two of whom had young children and two of whom did not. They found that when the women interacted with their dogs, PDS was produced by women with young children and by women who did not have young children. However, this study used a very small sample of participants and it is not clear from their report whether the two women who did not have young children at the time of testing had had children in the past. Nevertheless, this study makes progress in addressing the sampling issues of previous work, though further research is needed with wider populations in order to discover whether PDS is produced universally, or is simply an overgeneralisation of IDS in western culture, in which we often consider our pets to be part of the family.

While evidence so far suggests that humans alter their speech to accommodate different listener ability, it is not clear when sensitivity to linguistic competence develops or what skills are necessary to produce audience-appropriate speech. Batliner, Biersack and Steidl (2006) found that children used mostly child-directed and pet-directed speech when talking to a robotic dog. This suggests that the production of PDS occurs from an early age in both boys and girls. Similarly, Levin and Hunter (1982) studied the speech used by children aged between 5 and 12 while each played alone with a rabbit. They found that both boys and girls use PDS spontaneously. Levin and Hunter found no difference in the speech types used between age groups, however speech did differ between genders. Boys used simpler speech with more imperatives, while girls were more affectionate and used a larger range of pitch. The PDS used by the children (boys and girls) showed a similar pattern of acoustic properties to that of adult-produced IDS described in the literature, though child-produced IDS was not directly measured in this, or any previous study. If child-produced DDS is more similar to IDS than adult-produced DDS, this would suggest that young children may understand the affective needs of a listener, but perhaps at this stage do not understand the differences in linguistic potential. Levin and Hunter suggest that selection of the most socially appropriate speech type requires successful language development, but also integration of sophisticated sociocognitive skills. These may not be fully developed in younger children, which may give rise to overgeneralization of linguistically relevant features, including the hyperarticulation of vowels that is present in adult-produced IDS, but not DDS.

Recently, Ben-Aderet et al., (2017) found that DDS is produced in the absence of live dogs, when participants are presented with pictures of puppies, adult dogs, and senior dogs. In this experiment participants read scripted sentences to the pictures of dogs, and again to the adult experimenter. They found significant differences in the acoustic properties of DDS compared with adult-directed speech. Specifically, DDS was higher in pitch with more pitch variation over time and had higher harmonicity, confirming previous findings of Burnham et al (1998). The authors also showed that pitch was significantly higher for puppy-directed speech than for speech directed towards adult dogs or older dogs. By using pictures of dogs, Ben-Aderet et al., eliminated the issue of familiarity with the listener that was present in previous studies, and allowed the testing of DDS to multiple individual dogs. Although this confirms that DDS is produced even to unfamiliar dogs, the participants of this study were also all female. This does not therefore address the questions surrounding the prevalence of DDS production in wider populations including by men. It is also important to note that while this study indicates that DDS is produced in response to photos of dogs, with a similar pattern of results to that produced in response to real dogs, it is not clear from this study alone whether the DDS produced to the photos is representative of DDS produced to live models. It would be interesting to investigate whether there exist acoustic differences between speech directed at photos, compared to real listeners, especially since previous research has indicated that we are sensitive to the linguistic and affective needs of a listener, an aspect that clearly differs between real listeners and photos.

Although questions still remain as to how robustly DDS is produced across individuals of different genders, personalities and dog-experience levels, at least in Western societies DDS does seem to be produced widely. Why people address their dogs in this speech type and the function of this type of speech from the dog's perspective are explored in the next section.

Perception of speech

It is evident from the studies outlined above that speakers are sensitive to their audience in terms of acoustic preference, emotional needs and linguistic potential. It is important to note here that the production of audience specific speech is automatic, and, as far as can be determined, unconscious on the part of the speaker. In terms of changing the behaviour of the listener, there is evidence to suggest that human infants show a preference for IDS from a very early age (Kaplan et al, 1995). Cooper and Aslin (1990) measured preference for attending to IDS over ADS in 1-month- and 2- day-old infants. They found that both the 1-month-olds and the newborns preferred IDS over ADS, illustrated by longer looking durations towards IDS than towards ADS. This suggests that preference for IDS is present from birth and does not require post-natal experience. However, it is important to note that some aspects of language, including prosody, are audible in utero and therefore prenatal experience may also play a role. Further, Werker and McLeod (1989) measured affective responsiveness to ADS and IDS in 4-5 and 7-9 month old infants. Two trained raters judged the affective responsiveness of infants, comprising of how much they thought the infant was trying to interact with the speaker, how interested they appeared, and the valence of the infant's emotional state. They found that infants of both age groups showed greater affective responsiveness to IDS than to ADS. They also found that when presented with video recordings of infants listening to speech, unfamiliar observers rated the infants more ‘appealing’ when the infants were listening to IDS than when they were listening to ADS. This indicates that the use of IDS may facilitate the development of an emotional bond between adults and infants.

Conversely, although there have been several studies exploring the production of PDS by humans, there is little literature surrounding the effect of PDS on the behaviour of the animal to whom it is directed. It is possible that DDS is simply an overgeneralization of IDS in an attempt to bond with our pets, in a culture where pets, and dogs particularly, are often given the status of infants. If the main purpose of IDS is to facilitate language acquisition, and the modification with highest linguistic benefit (hyperarticulation) is missing from PDS, it is reasonable to suggest that the purpose of PDS is not to help our pets understand speech. Rather, supposing that PDS is functional, and not an overgeneralization of IDS, it is most likely that PDS is a more basic mechanism for gaining attention and strengthening the affiliative bond that humans share with their pets.

There are several possible reasons why dogs might show a preference for attending to DDS over ADS. It is possible that dogs learn over their individual lifespan that DDS is the type of speech their human uses when speaking to them, and through positive reinforcement individual dogs learn to pay attention when they hear DDS. Alternatively, domesticated dogs may have evolved to be sensitive towards DDS as it increased their chances of survival in a human-centred environment. Cooperation with humans is seen in dogs far more than in wolves, and it is possible that DDS facilitates cooperation with humans and functions for social benefit for both the human and the dog. As with IDS, it may also function to provide information about appropriate caregivers – indicating to the dog that the human is not a threat. This would have been important during the domestication process where living in a human dominated environment could have been dangerous. Future research recruiting dogs with no experience of DDS (such as newborn puppies), could reveal whether DDS is a learnt response in individuals, or if it is a more basic response to frequency or exaggerated intonation contours, that has evolved to promote survival and social bonding in domesticated dogs. It is important to note also that if preference to DDS is driven by a basic preference for pitch and intonation, it is possible that pet owners are positively reinforced by the dog's natural response to DDS, as it appears to increase attention and increase affiliative behaviour.

In order to establish whether dogs display any behavioural preference for DDS over ADS, Ben-Aderet et al. (2017) conducted playback experiments using the stimuli recorded to photos of dogs. They recruited 10 puppies and 10 adult dogs from a local shelter. During the playback experiment the dog was placed in a familiar room with a familiar human experimenter. A loudspeaker was placed in one corner of the room and the experimenter stood in the opposite corner. They selected approximately 30 seconds of dog-directed speech and 30 seconds of adult-directed speech and measured the dogs' behavioural preference to each type of speech with a composite measure that included latency to vocalize, total duration of looks towards the loudspeaker, duration of stays less than 1m from the loudspeaker, and duration of head tilts towards the loudspeaker. Puppies showed a higher behavioural response to DDS than to ADS, but this preference decreased as a function of age. Puppies also showed an increase in behavioural response with increased mean pitch, but this relationship was not evident in adult dogs. The authors conclude that puppies are highly reactive to DDS and that pitch is a key feature in modulating this preference, but that adult dogs do not react differentially to DDS and ADS. They argue that DDS may have a functional value in puppies, but not adult dogs, and therefore the use of DDS with adult dogs may simply be a 'spontaneous attempt to facilitate interactions with non-verbal listeners' (BenAderet et al., 2017, p1).

This is the first experiment to date that has investigated behavioural response to DDS and ADS in dogs. It very clearly shows that puppies are more sensitive to prosodic differences in DDS and ADS than adult dogs. However, there are several limitations of this work that should be addressed in order to determine whether their findings for adult dogs are robust. Firstly, it has not yet been verified that DDS produced to pictures does not differ from those produced by live models. It is possible that the speech stimuli used in this study were not ecologically valid. It may be that puppies are less selective in the type of DDS they respond to, whereas for older dogs, with more experience of DDS, the DDS in this study did not differ sufficiently from ADS to elicit a differential response. Secondly, the same script was used for DDS and ADS, which was highly dog relevant (e.g. 'who's a good boy? Come here! Good Boy!'). It is also possible that the puppies, with little knowledge of human language, rely wholly on prosodic features to identify relevant speech. The adult dogs may respond to the dog-related content in ADS, such that the behavioural response to DDS was not significantly higher. A follow-up study where dog- and adult-relevant content is controlled, could identify whether prosody, or content alone is sufficient for driving a preference for DDS in puppies and adult dogs. Finally, in this study, speech was played from a speaker in the corner of the room where no human was present. If preference for DDS functions as a way of improving social communication, it may only be relevant to display this preference if there is a human with which to socialise. It is possible that if no human experimenter is present, adult dogs realise that there is no social benefit to reacting preferentially to any speech. Puppies, with little experience of the world, may not recognise this, and therefore still responded to DDS in the absence of a feasible producer.

It is clear that several questions remain about the behavioural response by dogs to DDS, including their sensitivity to content. In reality, communication with our pets rarely occurs in the absence of other cues such as facial expressions, eye-gaze and body language indicators. We already know that dogs are sensitive to these cues (Hare, Brown, & Tomassello, 2002; Miklosi et al, 2003; Hare et al., 2005), and so it would be interesting to explore how preference for DDS may interact with these behaviours. Although to date I have focussed on human-dog communication, dogs also produce and receive signals in a range of modalities from each other. What dogs understand from the auditory signals produced by conspecifics will be the focus of the next section.

Communicating with Conspecifics

Wolves, while capable of producing many vocalizations, in reality produce a far narrower range of vocalizations than domestic dogs (Bradshaw & Nott, 1995; Cohen & Fox, 1976; Fox, 1971). Wolves will vocalize in very specific contexts, such as defence of territory, guarding of food, and recruitment of group members, but rely on growling, and howling sounds, rather than barking for the most part (Harrington & Mech, 1979; Klinghammer & Laidlow, 1979; Theberge & Falls, 1967; Joslin, 1967). In contrast, domestic dogs use vocalizations in a wide range of contexts, including greeting, play, submission, defence, and being left alone (Cohen & Fox, 1967). It has been suggested that the variety of barks and the contexts in which they are used, are less specific than vocalizations in wolves and other canids, and therefore that vocalizations by dogs are less communicative, and instead may serve to simply attract the attention of the receiver (Coppinger & Feinstein, 1991).

In the wild, vocalizations are useful for enhancing physical and olfactory communicative displays, and are able to replace these signals for example across large distances, or at night (Cohen, & Fox, 1976). These communications vary widely within canids, and interspecies comparisons of vocalizations and the contexts in which they occur can help us to understand whether barking in domestic dogs has emerged as a non-functional by-product of domestication, or whether it still retains communicative information. Generally speaking, there are two developmental stages of vocalization in canids. At birth, vocalizations include yelping, whining, and screaming, and are thought to decrease distance between the infant and the mother (approach type calls). Later to develop are vocalizations that elicit withdrawal of conspecifics, to create distance. These include growls, clicks, and barks, and appear to emerge as the animal becomes more independent. Cohen and Fox (1976) recorded vocalizations from a variety of species hand-reared in captivity including wolves, foxes (red and grey), coyotes, arctic foxes, jackals, and domestic dogs, as well as zoo animals including dingos, New Guinea singing dogs, maned wolves, culpeo, bush dogs, and cape hunting dogs. They recorded calls from these species at various ages (birth to ~3 years) and in different contexts. Vulpine (fox-like) and canine (dog-like) groups demonstrated largely distinct repertoires, with vulpine species showing high incidence of coo-calls, clicks, yips and screams, and an absence of group vocalizations, in contrast to canines. High incidence of barks in many different contexts was exclusive to domestic dogs. Domestic dogs showed later development of second stage withdrawal barks, remaining longer in the juvenile stage. Infantile sounds such as yelps and whines persist into adulthood of domestic dogs, consistent with the domestication syndrome described in the self-domestication hypothesis.

Some argue that this increase in barking in domestic dogs is a non-functional by-product of domestication (Coppinger & Feinstein, 1991). While increased frequency of barking may have arisen as part of the domestication syndrome, it is also reasonable to suggest that the selection pressure for silence and stealth in wild animals is relaxed in domestic dogs, who are able to enjoy the increased safety and protection that comes with living in human groups. The relaxing of this selective pressure may also explain the increased incidence of barking in dogs, and may have led to the barks of domestic dogs losing the context specificity that is seen in the vocalizations of wolves. Context specific vocalizations can sometimes function referentially, in that they provide the listener with information about the eliciting event that enable them to respond in a similar way to the call as they do to the original event that elicited the call (Evans et al., 1993). To determine whether dogs extract meaningful information from conspecific barks, we must investigate whether they can use barks of another dog to inform their behaviour in an ecologically valid setting. The seminal investigation into the informational content of animal calls used playback experiments to probe the understanding of alarm calls in vervet monkeys (Seyfarth, Cheney & Marler, 1980). These monkeys give acoustically distinct alarm calls depending on the type of predator approaching, and the listeners' understanding of these different alarm calls is demonstrated by their contextually appropriate behavioural responses to playbacks of these calls. For example, vervet monkeys produce acoustically different alarm calls in the presence of leopards, eagles, and snakes. In response to playbacks of leopard alarm calls, conspecifics move into trees, while upon hearing an eagle alarm call they look up, and snake alarm calls result in scanning of the ground. Recordings of each alarm call were played in the absence of any predator, to isolate the information individuals could extract from the calls alone without the other cues that normally accompany a predator approach. The monkeys' predator-specific behaviour in response to playbacks provides strong evidence that vervet monkeys use acoustic properties of conspecific alarm calls to extract information about the likely eliciting event and use this information to inform their subsequent behaviours. This pioneering study also established the use of playback experiments for obtaining meaningful behavioural measures that give insight into understanding in non-linguistic subjects.

Playback experiments have been successfully used to probe listener understanding in a wide variety of species, including dogs. Dog understanding of growls produced in play, stranger approach and food-guarding contexts was investigated by Farago, Pongracz, Range, Viranyi, and Miklosi (2010). Dogs were less likely to take a seemingly unattended bone following playback of a growl recorded in a food-guarding situation, than for play growls and growls recorded during the approach of a stranger. It is interesting that the two agonistic growls led to significant behavioural differences despite largely similar acoustic properties, showing that dogs are processing something more than simply positive or negative valence in the growls.

While it seems that dogs are able to extract information from the growls of conspecifics, it is perhaps more surprising that the contextual barks of domestic dogs, may also provide information cross-species, with humans. In 2005, Pongracz, Molnar, Miklosi & Csanyi recorded barks from a Hungarian herding dog breed (Mudi) in various different situations including *Alone, Stranger, Play, Ball* (hold ball or favourite toy out of reach), and *Schutzhund* (trainer encouraged dog to bark aggressively and bite his bandaged arm). These recordings were played back to humans with varying degrees of experience with dogs, and with Mudis in particular. The authors found that the human participants were able to identify the contexts of these barks significantly above chance, regardless of their levels of experience. Interestingly, when asked about the affective state of the dog during the playbacks, human ratings of the barks corresponded with the situation in which the bark was produced (e.g. Play rated higher in happiness and playfulness, Schultzhund rated higher in aggressiveness). This suggests that barks not only convey contextual information, but that humans are also able to extract emotional information about the producer, from the acoustic properties of the bark, and challenges the idea that increased vocalizations, specifically barking, in domestic dogs is a non-functional by-product of domestication. Instead it appears that vocalizations in domestic dogs may contain contextual and emotional information that may also aid communication cross-species, with humans. Perhaps surprisingly, in contrast to research effort to understand cross-species understanding of barks, we know very little about how dogs perceive each other's barks; something that I attempt to address in Chapter 5.

Thesis Outline

The domestic dog shares a special relationship with humans, and provides a unique opportunity to study the evolution of social communication in a cross-species model. This is particularly true, as we are able to compare the social and communicative skills of dogs, with their closest living relative and model of their last common ancestor, the wolf. The research presented in this thesis focuses on understanding the ability of dogs and humans to communicate with each other as well as investigating communication between dogs.

Chapter 2. Individual Differences in the Production of Dog-Directed Speech. This empirical chapter examines the acoustic properties of speech used to communicate with pet dogs, with a focus on individual differences that may influence the production of dog-directed speech, including gender, dog experience, and empathetic abilities.

Chapter 3. Dogs prefer naturalistic dog-directed speech. This empirical chapter investigates attentive and affiliative preference displayed by dogs for dog-directed speech compared with adult-directed speech. Factors influencing this preference are examined, including speech prosody, and speech content.

Chapter 4. The interaction of speech type and eye-gaze in communications with domestic dogs. This empirical chapter aims to examine whether dogs display a preference for direct or averted eye-gaze, and the interaction of eye-gaze with speech type.

Chapter 5. Contextual Barking: A cross-species cue for social communication. This empirical chapter investigates whether domestic dogs are able to use the contextual information in barks of conspecifics, to alter their behaviour appropriately.

Chapter 6. General Discussion. This chapter brings together results from chapters 2-5 in order to discuss how these studies have contributed to our understanding of social communication in domestic dogs. I will also discuss possible applications, and directions for future research.

Chapter 2: Individual Differences in the Production of Dog-Directed Speech

Abstract

Previous study of infant-directed speech (IDS) and dog-directed speech (DDS) has shown that these special speech types demonstrate higher pitch, and pitch modulation compared with adult-directed speech (ADS), and appear to be produced automatically and unconsciously when talking to listeners to meet their linguistic and affective needs. However, while it is often argued that IDS and DDS are universally produced, most investigations measure speech of females, particularly mothers, and very little has been considered in the way of individual differences in their production. I aimed to investigate the acoustic properties of DDS produced by males and females in the presence of an unfamiliar dog, and also examined the influence of dog-ownership and levels of empathy on speech. The results suggest that both males and females produce DDS which is higher in pitch and pitch modulation than their ADS, though females increase their pitch for DDS to a greater extent than males. No significant effects of gender were found on pitch modulation, nor on the content of interactions. There was also no difference in the acoustic changes or the content of DDS compared to ADS, between dog-owners and non-owners, and no significant influence of empathy. These results also found no evidence of hyperarticulation of vowel sounds for DDS, which is in line with previous findings and suggests that hyperarticulation may be the aspect of IDS that aids spoken language acquisition. Overall, these results highlight that the speech of males and females to infants and dogs may contain acoustic differences, and that therefore the production of these characteristic changes in IDS and DDS may not be as universal as previously described.

Introduction

It has been known for some time that adults use a special speech type when talking to human infants. This speech, known as infant-directed speech (IDS) is characterised by increased pitch, greater pitch modulation, increased emotional affect, and hyperarticulation of vowel sounds (Burnham, Kitamura & Vollmer-Conna, 2002). IDS allows infants to select appropriate social partners (Schachner et al, 2011), and increases social bonding between an infant and its caregivers (Kaplan et al, 1996). However, it has been argued that the main function of infant-directed speech is to facilitate language acquisition, through the amplification of the phonetic characteristics of native language vowels (Kuhl et al, 1997). When speaking to their infants, mothers do not simply raise the frequency of all speech formants, as would be the case if mimicking childlike speech. Instead, the difference between the first and second formants appears to be selectively greater in IDS compared with adult-directed speech (ADS). This is reflected acoustically in exaggerated /i/, /a/, and /u/ vowel sounds, and was originally thought to improve the clarity of speech. This hyperarticulation of vowel sounds also occurs when speaking to a foreign adult, or second-language learner, who may have similar linguistic needs to those of a human infant, but these listeners are unlikely to possess the same affective needs (Uther, Knoll, & Burnham, 2007). Uther et al, found that hyperarticulation is similar for IDS and foreigner-directed speech but that affect is lower in foreigner-directed speech than even ADS. This may suggest that acoustic modifications such as hyperarticulation are pedagogical and occur independently of changes in pitch, and affect. From this it is reasonable to argue that while the changes in pitch, and affect of different speech may facilitate social relationships, it is the hyperarticulation of vowel sounds that aids in language acquisition. Moreover, recent research by Martin et al, (2015) has found that the phonetic distribution of the vowel sounds /i/, /u/, and /a/ overlap more in IDS than in ADS, leading to poorer performance of speech recognition algorithms. This suggests that the mechanism by which hyperarticulation facilitates language acquisition may be more complex than simply improving clarity.

While the mechanisms underpinning IDS's beneficial effects on language learning and social bonding are largely unknown, it is thought that IDS is a robust-cross-cultural phenomenon. There is evidence for IDS across several languages, including English, Japanese, Swedish, French, Russian, and German (Fernald et al., 1989). For instance, Fisher and Tokura (1996) studied IDS produced by English and Japanese mother, and were able to identify exaggerated intonation contours, vowel lengthening, and local acoustic changes in duration and pitch across the two languages. It is clear however, that the vast majority of studies have occurred in WEIRD (western, industrialized, educated, rich, and democratic) populations. For instance, there is very little systematic investigation of the properties of IDS in more traditional, hunter gatherer or rural subsistence third-world societies. One recent study by Broesch and Bryant (2015) examined pitch and speech rate of mothers speaking to infants and adults in Fiji, Kenya, and North America. They found that in all three cultures, mothers used higher pitch when talking to infants compared to adults, and used a greater range of fundamental frequencies. Interestingly, previous studies have claimed that American mothers produce greater pitch modulation than mothers in other cultures, but Broesch and Bryant found that this was not the case after controlling for maternal education. This suggests that maternal education may be another factor influencing the production of IDS, and highlights that education and socioeconomic background should also be considered when assessing the universality of IDS. This study begins to fill the gap in the literature with regards to cultural differences in IDS, suggesting that there are similar patterns of acoustic changes in IDS produced in Western and traditional societies, but does not address that the majority of studies that discuss the acoustic properties of IDS, do so with female speech. It is possible that this is due to recruitment of mothers with infants, but there is far less evidence investigating whether men spontaneously produce IDS, and whether the acoustic properties differ between genders. One study by Fernald et al., (1989) acoustically analysed the speech of mothers and fathers in French, Italian, German, Japanese, British English, and American English. They found that mothers and fathers increased the pitch of their speech for infants compared with ADS, and highlight that mothers, but not fathers, use a wider range of frequencies in IDS compared to ADS. They also found that pitch modulation was greatest in American English compared to other languages. This study suggests that while IDS may be produced in several languages, the acoustic properties of IDS in each language may not be as universal as previously suggested. The difference in speech by males and females also highlights the need for more extensive study of individual variation in the production of IDS.

As discussed in Chapter 1, Pet-directed speech shares some of the acoustic properties of IDS, including high frequency and exaggerated affect (Burnham et al, 2002), but does not contain hyperarticulation of vowel sounds thought to aid in language acquisition (Burnham et al, 1998; Kuhl et al, 1997). There is also evidence that hyperarticulation may be modulated by the perceived linguistic capacity of the receiver (Xu, Burnham, Kitamura & Vollmer-Conna, 2015). Xu et al, found that while participants hyperarticulate their vowels most when talking to infants, hyperarticulation also occurs when speaking to parrots, but not to dogs. Parrots are often perceived to be able to learn language which is in contrast to dogs who are unable to produce speech of any kind. These studies suggest that we are able to unconsciously alter the acoustic properties of speech depending on the linguistic potential of the listener, as well as their social and affective needs.

Although more recently the focus of this literature has been on acoustic properties of IDS and DDS, earlier studies focused mainly on the content of IDS and DDS interactions. To do this, early work relied mainly on parental or observer reports of the content of speech (Bynon, 1968; Ferguson, 1964; Voegelin & Robinett, 1954). These studies found that IDS interactions contained phonological simplification, repetition, and frequent addition of suffixes to words (such as dog becoming doggy, or horse becoming horsey; Kaye, 1980; Phillips, 1973). They also found that compared with ADS, IDS showed a decrease in type/token ratio, which is an indicator that mothers reduce the number of unique words used when talking to their infant, compared to the number of unique words one would expect to hear in ADS (Henning, Striano & Lieven, 2005; Remick et al, 1976). Further, the lengths of utterances in IDS tend to be far shorter than ADS, with reduced words per utterance (Phillips, 1973; Snow, 1977).

There are some similarities between the lexical structure typical of IDS and that of DDS, including short utterances with high repetition. Mitchell and Edmonson (1999), found that participants used short, repetitive utterances, with 7 words (including come, the dog's name, and ball) accounting for more than half the words used, and commands (imperatives) accounting for over 50% of utterances. They also found that approximately 10% of utterances were questions. Mitchell (2001) supported these findings, and also showed that both IDS and DDS included a high number of *attention-getting devices* such as clicks, claps, and non-verbal exclamations (e.g. gasps). Some differences between IDS and DDS included that DDS contained shorter sentences and more repetitions, where IDS contained a higher number of questions, declaratives, and deictic utterances (this, that, these, you etc). These findings demonstrate that speakers tend to simplify their speech, and include more surprising, attentive and affiliative content for infants and dogs, as well as modulating the acoustic properties of their speech. Importantly, all of these studies are conducted with female participants, who are also mothers, and whether these findings also extend to other populations including males, and non-mothers, is an outstanding question.

Until recently there has been little investigation into individual differences in the production of IDS and DDS. As a result, it is not yet clear whether the acoustic properties of IDS and DDS revealed in previous studies are unique to females, or to mothers more specifically, due to the bias in previous samples. One study observed 10 male and 15 female dog owners during free interactions with their pets (Prato-Previde, Fallani & Valsecchi, 2006). They found that men and women differed in their use of verbal communication, where women talked more than men, and showed lower latency to start talking at the start of the interaction, and after periods of quiet. The authors observed that for females, the acoustic properties of their DDS more closely resembled IDS, while participants reported similar levels of attachment, and displayed similar levels of affiliative play. They conclude that the differences in DDS displayed by males and females are in concordance with an evolutionary scenario whereby women are pre-disposed to use language as a communicative tool. This study is interesting in highlighting the potential differences between the tendencies of men and women to interact vocally with an infant or a pet, but as they do not conduct any systematic analysis of the acoustic features, it is not clear whether the speech itself used with the dogs differed between males and females. A study by Mitchell (2001) also found that females were more likely to use *DDS* when talking to a dog, compared with males, but interestingly, Mitchell found that both men and women used this speech more when interacting with an unfamiliar dog, compared with a familiar one. This further highlights the need to conduct studies with more representative samples, as it appears that gender and familiarity may have more influence on the production of infant- and dog-directed speech than previous studies would suggest, which questions the claim that these special speech types are universally produced.

Further, as it has been shown that IDS and DDS are important for affiliative bonding, it is reasonable to suggest that the type of speech produced may be influenced by the ability of the speaker to understand the emotional needs of the infant. For example, individuals who find it difficult to understand others, such as those with autism spectrum disorder, may not produce acoustically appropriate speech types in the presence of infants and dogs (although this remains empirically untested). Difficulties in social communication in autism spectrum disorders have been linked to dysfunction in mechanisms involved in empathy (Dapretto et al, 2006) and it would be interesting to investigate whether individual differences in an individual's ability to empathise with a listener, influence the production of infant- and pet-directed speech. It is possible, however, that experience also influences the ability to empathise with the emotional needs of others. For example, it has previously been shown that mothers are more accurate than non-mothers in identifying different infant cries, and that they become more accurate with increased experience after birth (Sagi, 1981). The latter suggests that while empathy is important for understanding the needs of others, it is likely that experience also plays a role. Thus it may be reasonable to expect that both empathetic ability and experience with the target audience (e.g. infants or dogs) may influence speech production, but these aspects of individual differences have not been examined previously in relation to the production of IDS or DDS.

The following study aimed to investigate the individual differences in the production of dog-directed speech. Spontaneous naturalistic speech produced upon meeting an unfamiliar dog for the first time was recorded. I explored whether individual differences including gender, dog-experience, and empathy influence the acoustic properties of DDS. Based on previous research, I hypothesised that DDS would be higher in pitch than ADS, and would have greater pitch modulation. I also expected based on the results from Burnham et al (2002) that DDS would not show the hyper-articulation of vowel sounds that is observed in IDS. I expected that the difference between acoustic features (e.g. pitch, pitch modulation) in an individual's ADS and DDS would be greater for females than males. I also hypothesised that there would be an effect of experience with dogs, with dog owners expected to produce a greater increase in pitch and pitch modulation for DDS compared to non dog-owners. Finally, given that DDS partly serves to address emotional needs of the listener, I hypothesise there will be an effect of empathy on the acoustic properties of DDS, where participants with higher empathy scores would produce greater differences in pitch and pitch modulation for DDS compared with their ADS than those with lower empathy scores.

Methods

Participants and study site

Participants were 30 students from the University of York, recruited through opportunity sampling. Fifteen males and 15 females took part (mean age = 21.7 ± 2.3 years).

Materials

Speech was recorded as uncompressed wav files using a Marantz PMD661 solid-state recorder set to sample at 96 kHz. Three children's toys, a shark, a shoe, and a sheep were provided, about which the participant could talk to the dog. This was to produce vowel sounds required for analysis of vowel hyperarticulation, and the toys chosen were based on those used in Burnham et al (2002). A questionnaire was administered in order to evaluate empathetic abilities of the participants. The questions used were those from the Questionnaire of Cognitive and Affective Empathy (QCAE), where scores were summed for analysis as described by Renate et al., (2011). Participants also completed a questionnaire asking their gender and whether they were a dog owner or not. The dog in this experiment was a 3 year old female Parson-Russell Terrier, named Lily, who was tethered in the experimental room with access to a bed and water. A line was drawn on the ground half way between the door and the dog, which participants were instructed to stay behind during the first part of the experiment, and when presenting Lily with the toys.

Design

In a between-subjects design, mean differences in pitch and pitch modulation between DDS and ADS were compared for males vs females, and for dog owners vs non-owners, as well as across the range of empathy. Hyperarticulation of vowels in DDS and ADS was compared in a within-subjects design.

All participants recorded ADS1 before the trial in a room adjacent to where the dog was kept. After the trial they recorded ADS2and the ADS recording used for analysis was counterbalanced, with either ADS1 or ADS2 recordings = randomly selected for analysis (ADS1 N=15, ADS2 N=15).

Procedure

Participants entered the room adjacent to where the dog was kept for the pre-trial interview. In this room participants were informed about the procedure before signing a consent form and risk assessment, as required by the ethics committee, and the participants completed the questionnaires. The experimenter then started the sound recorder and asked participants questions to obtain ADS1. Once a few minutes of ADS was recorded, the experimenter pointed to the toys in turn and asked participants to name the toys. If the participant called the toy by a different name (e.g. Lamb), they were corrected and asked to name the toy again to obtain ADS recordings of the toy names for vowel analysis (Sheep, Shark, Shoe). Then participants were informed that in the next room there was a friendly dog called Lily who they would be able to interact with. They were informed that if at any point they felt uncomfortable they could leave the room. Participants were then brought into the next room where the dog waited with a second experimenter. The sound recorder was started before the participant entered, and they were instructed to interact with the dog from behind the line for approximately 1 minute. The experimenter wore headphones and did not interact with the participant other than to read the standardized instructions for the remainder of the trial, to avoid influencing their speech. Participants were instructed to ‘interact with lily from behind the line.’ after 30 seconds to 1 minute participants were told, ‘you may cross over the line and interact with Lily physically if you wish’, only if the dog appeared at ease with them. After this time (30 seconds – 1 minute) participants were asked to ‘step behind the line and name the three toys’. They were then instructed to ‘talk to Lily about her toys, using the names discussed, but please do not cross the line or give her the toys’. They were given approximately 1 minute to discuss the toys, and if they had not used all the names after 30 seconds, the experimenter told them to please talk about all three toys using their names. After this time, the toys were removed and the participant was taken back into room one. In room one the sound recorder was started and the participants were asked some questions about the experiment, and to name the three toys again. Finally, participants were debriefed and asked to wash their hands before leaving the building. This procedure provided recordings of ADS1, DDS, and ADS2, with names of toys for acoustic analysis.

Acoustic Analysis

ADS and DDS speech recordings were analysed in PRAAT (version 6.0.35) with pitch settings 75-500Hz. Pitch was measured using cross-correlation which compares the similarity of the signal with a time-displaced signal and is thought to be optimal for voice analysis. Speech segments with a continuous pitch line were measured in terms of mean F0, min F0, and max F0 using the *get pitch* function. Speech was analysed from the beginning of each interaction (ADS or DDS), until at least 5 seconds of pitch-tracked speech was measured (Mean duration measured ADS = 6.74 ± 3.00 seconds, DDS = 4.96 ± 3.17 seconds). For ADS segments, this usually occurred in the first 10-20 seconds of speech, compared to DDS segments, where utterances were usually shorter and further apart, and therefore longer segments were measured. Pitch modulation was calculated as maxF0-minF0. Mean pitch, and pitch modulation were calculated for ADS and DDS of each participant, and the difference between these (DDS-ADS) was obtained to represent the change in pitch and pitch modulation between the two speech types.

For the words Sheep, Shark, and Shoe, the first (F1) and second formants (F2) were measured using the *get formants* function in PRAAT (Max formant 5500Hz, Number of formants extracted: 5, Window length 0.025s, Dynamic Range: 30dB, Dot size: 1). Only the vowel sounds /i/ /a/ /u/ were extracted from the centre of the words sheep, shark, and shoe. F1, and F2 for these three vowel sounds of DDS and ADS were measured.

Transcripts were produced for each recording (see appendix 1), and the content was evaluated for the first 30 seconds from the start of the first utterance. Number of words spoken, and number of utterances were counted (with each sentence counting as one utterance), and mean length of utterance was calculated (number of words/number of utterances). Also counted were the number of exact repetitions where counted as the number of times each word or phrase is said during the first 30 seconds of the recording. Repetitions did not need to be consecutive ('hi' 'hi' 'hi' = 3, 'come here' 'you're good' 'come here' = 2). The number of questions, greetings (hello, hi), and imperatives (commands such as come here) were counted, as well as the number of attention getting devices, such as clicking fingers/tongue, gasp/sharp intake of breath, kissy noises, tapping floor or body, and clapping hands.

Inter Observer Reliability

All recordings were analysed by the primary coder (AB). A second coder blind to the hypotheses of the experiment analysed the mean, minimum, and maximum pitch of ADS and DDS recordings from 10 participants (N=30/90 measures). The primary observer had high agreement with the secondary coder across these measures (Spearman's R = 0.94, *p*<.001). For vowel hyperarticulation, the secondary coder measured the first and second formants of the vowels in the words *Sheep, Shark*, and *Shoe* for 6/30 participants (N= 36/180 measurements). The primary and secondary coders had high agreement (Spearman's R = 0.91, *p*<.001).

Statistical Analysis

All data were analysed using IBM SPSS (version 24) with the alpha level set for initial analyses at p < .050. To assess how pitch characteristics in DDS and ADS varied between individuals, generalized linear mixed models (GLMMs) were conducted separately for the dependent variables Mean Pitch Change, and Pitch Modulation Change. The data for mean change in pitch met the assumptions for statistical testing using GLMMs, as demonstrated by the normality of the residuals (S-W (29) = 0.97, *p* =.542). The data for mean change in pitch modulation failed the assumption of normality of the residuals (S-W (29) = 0.85, *p* =.001), and therefore a log transform was applied to the raw data. Following transformation, the residuals met the assumption of normality (S-W (29) = 0.98, *p* = .918) and therefore the GLMMs for pitch modulation were conducted using these transformed values. Although ideally I would have entered all fixed factors and the interactions between them into single model, with only 30 participants, I did not have sufficient sample size to include all three fixed factors and four possible interaction terms in a single model. I therefore conducted separate GLMMs for each fixed factor *Gender, Dog Experience,* and *Empathy Scale,* using a normal probability distribution and identity link function. To account for multiple testing of the same dependent variable, I applied a Bonferroni correction to the alpha level for these models, resulting in an alpha of *p*<.017.

I then conducted post-hoc analyses of the interactions between the fixed factor which was highlighted as explaining a significant amount of variation in the data in the initial models, and the other fixed factors. To avoid overfitting the model, I did not include interactions that were not likely to explain variance in the data given the results from the previous models; this resulted in three GLMMs with fixed factors *gender*experience, gender*empathy,* and *gender*experience*empathy*, with a Bonferroni corrected alpha of $p<.017$ to correct for multiple comparisons.

To assess how content of DDS and ADS varied between individuals, a similar approach was taken. The raw data for number of words (S-W (27) = 0.93, $p = .067$), and mean length of utterance (S-W (29) = 0.96, $p = .381$) demonstrated normality of residuals, however number of utterances was negatively skewed and failed the assumption of normality of residuals ((S-W (29) = 0.89, $p = .007$) and therefore was transformed (log10). The transformed values for number of utterances was found to meet the assumption of normality of residuals (S-W (25) = 0.94, $p = .148$) so these transformed data were entered into the analyses. Separate GLMMs were conducted for the change in dependant variables (DDS-ADS), number of words, number of utterances, and mean length of utterances, using a normal probability distribution and identity link function. Three models for each dependent variable were run, each with one of gender, dog experience and empathy score entered as a fixed factor.

Through examination of the raw data, and assessing the distribution of the residuals, the remaining measures (imperatives, questions, exact phrase repetitions, attention getting devices, and greetings) failed to meet the requirements for inferential statistical testing (data not normally distributed even after transformation). Instead, descriptive statistics are presented for these measures.

Results

For all analyses, one male participant was removed as he only produced 4 words of DDS, three of which were the names of the toys, resulting in insufficient data from which to calculate average pitch changes (N=29).

As previous research has suggested that DDS is higher in pitch, and contains greater modulation in pitch compared with ADS, I first investigated whether DDS produced by each participant met these criteria for DDS. Each participant was deemed to have produced DDS if the mean pitch and pitch modulation for speech directed to the dog fell outside the 95% confidence interval for pitch and pitch modulation in ADS. Figure 2.1 demonstrates that 86.67% of females produced DDS that met the criteria for both pitch and pitch modulation. The remaining 13.33% of females produced DDS that met the criteria for pitch only. For males, however, only 57.14% of DDS met the criteria for both pitch and pitch modulation, and 7% of participants met criteria for neither pitch nor pitch modulation.

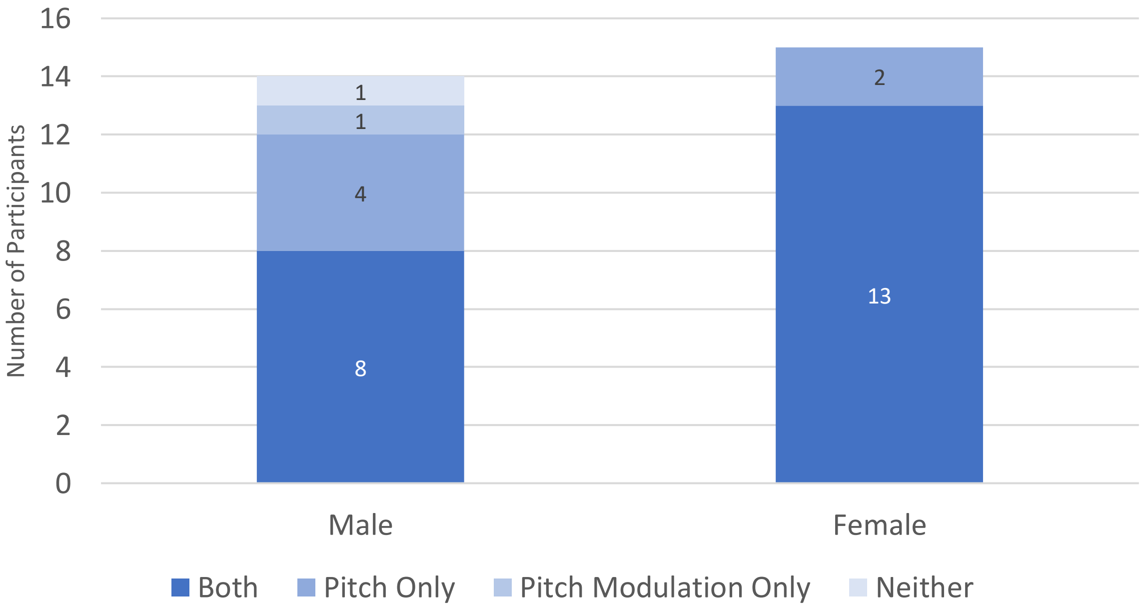


Figure 2.1. A stacked bar graph demonstrating the number of participants who met the criteria for DDS in terms of pitch and pitch modulation.

For each participant, changes in average pitch and pitch modulation between ADS and DDS were calculated (Mean DDS – Mean ADS) and are displayed in figure 2.2 below. The first GLMM revealed a significant main effect of gender on pitch (F (1, 27) = 7.70, *p* = .010), with females showing a greater change in pitch between DDS and ADS compared to males (see Figure 2.2). Figure 2.2 indicates that females show a greater change in pitch modulation than males, however a GLMM revealed that this was non-significant (F (1, 27) = 3.30, *p* = .082, Figure 2.2).

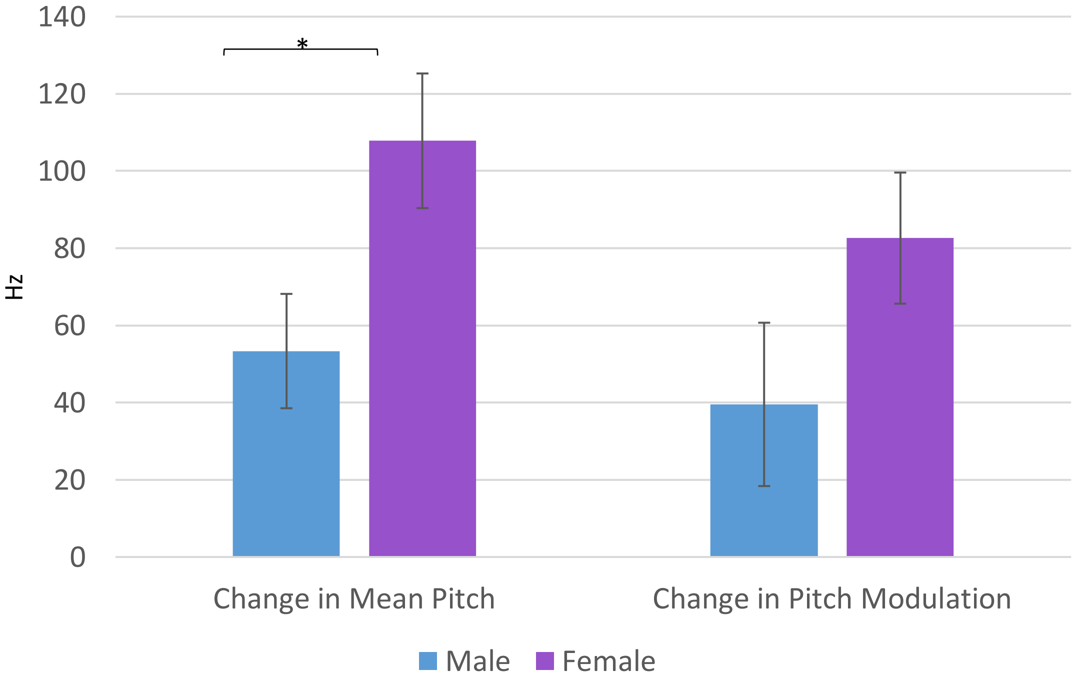


Figure 2.2. A graph to show the mean change in pitch and pitch modulation between DDS and ADS for males and females. Change was calculated as DDS-ADS, thus the positive values in the graph illustrate that on average both males and females produced higher pitch speech and greater pitch modulation in DDS than ADS. A significant main effect of gender was found for mean pitch (denotes p<.050) but there was no significant main effect of gender for pitch modulation. Error bars represent 1 Standard Error of the Mean.*

Two further GLMMs revealed no significant main effects of empathy or dog experience on pitch or pitch modulation (Table 2.1). There were also no significant interactions between these variables and gender.

Table 2.1. Results of GLMMs (1,27) for the influence of empathy and dog experience on pitch and pitch modulation, as well as the interactions between these factors and gender.

	Pitch **F*(p)***	**Pitch Modulation** **F*(p)***
Empathy	0.34 *(.562)*	0.66 *(.424)*
Gender * Empathy	0.04 *(.838)*	2.29 *(.144)*
Dog Experience	0.97 *(.334)*	3.08 *(.091)*
Gender * Dog Experience	0.55 *(.467)*	3.46 *(.076)*

Finally, hyperarticulation of vowels was measured by comparison of the first, and second formants for the vowel sounds /i/ /a/ /u/ from the centre of the words *Sheep, Shark, and Shoe*. These vowels are plotted below in the form of hyperarticulation triangles (f1-f2; figure 2.3). Vowel hyperarticulation triangles were plotted for each participant and the area of these triangles were calculated from the f1 and f2 coordinates for the three vowel sounds. The areas of DDS and ADS hyperarticulation triangles were compared using a paired samples t-test which found no significant main effect of speech type on vowel space areas (DDS mean =18266Hz^2, ADS mean = 19352Hz^2; t(28) = 0.67, p = .879).

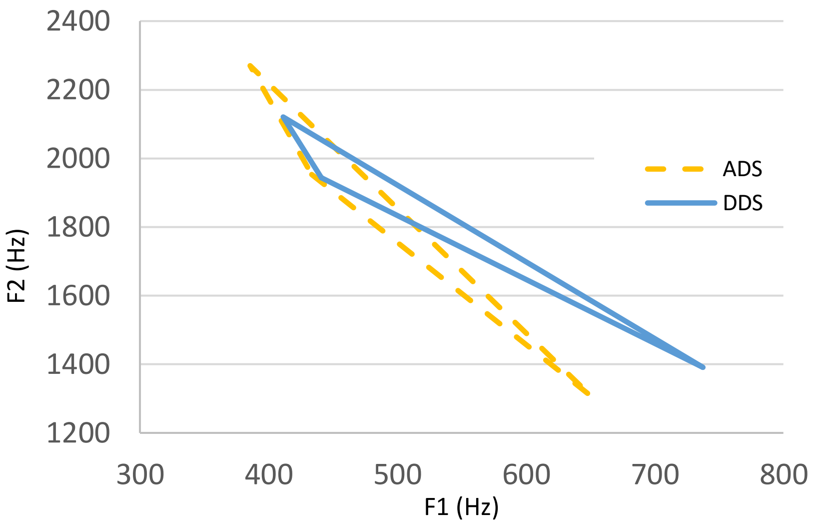


Figure 2.3. Hyperarticulation triangles plotted in f1-f2 space for dog- and adult-directed speech.

Semantic Content of DDS

For each participant, the mean length of utterances and the number of words, utterances, imperatives, questions, exact repetitions, attention getting devices and greetings were counted during the 30 seconds from the beginning of the first utterance. The number of words spoken and the mean length of each utterance were lower for DDS than ADS, while all other measures were higher for DDS than ADS. Changes (DDS-ADS) in number of words, number of utterances (log transformed), and mean length of utterance between the two speech types (DDS-ADS) were examined using separate GLMMs. GLMMs revealed no significant effects of gender, dog ownership, or empathy score on any of these three content measures (Table 2.2; Figure 2.4; Figure 2.5).

Table 2.2. Results of GLMMs for gender, dog ownership and empathy score, for measures of number of words, number of utterances (log transformed) and mean length of utterances.

	df	**Gender F (*p*)**	**Dog-Ownership F (*p*)**	**Empathy F (*p*)**
Number of words	1, 27	2.33 *(1.39)*	0.01 *(.928)*	0.02 *(.882)*
Number of Utterances	1, 23	2.35 *(.139)*	0.33 *(572)*	2.19 *(.153)*
Mean length of Utterance	1, 27	0.02 *(896)*	1.75 *(.197)*	1.38 *(.251)*

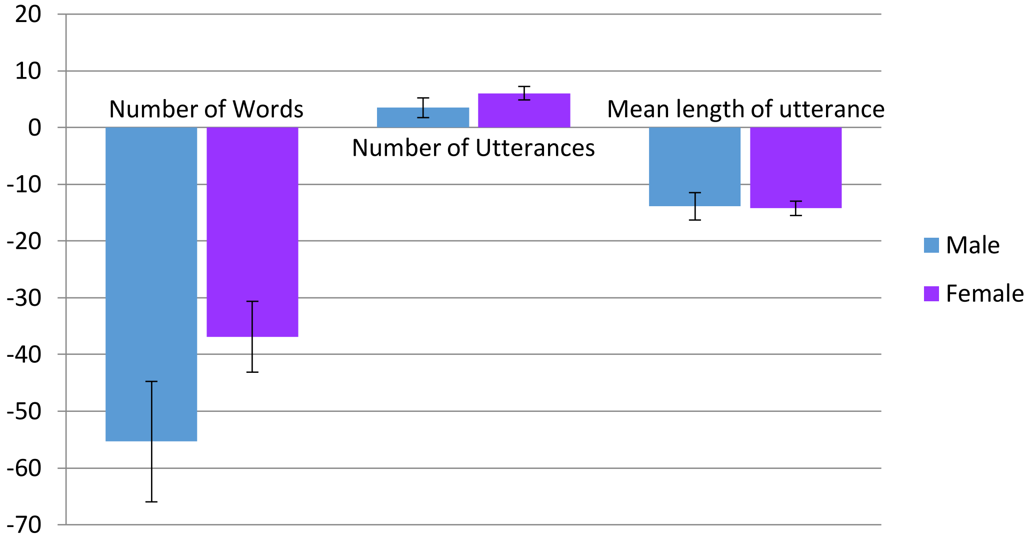


Figure 2.4 Mean change in number of words, number of utterances and length of utterances between DDS and ADS for males and females. Change was calculated as DDS-ADS, thus the positive values in the graph illustrate that on average both males and females produced more utterances in DDS than ADS and the negative values illustrate that on average both males and females produced fewer words and shorter utterances in DDS than ADS. Error bars represent standard error of the mean. GLMMs revealed no effect of gender for any of the three measures.

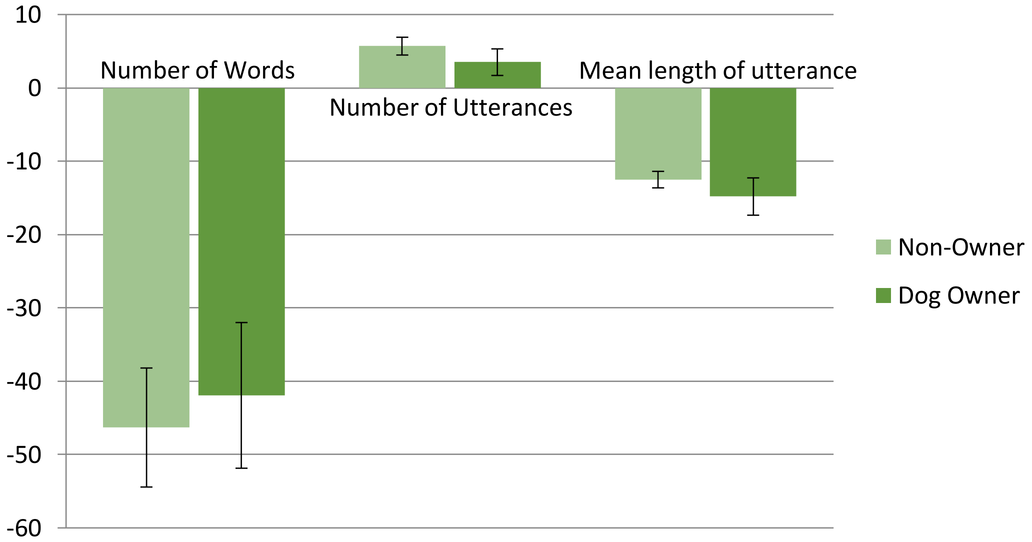


Figure 2.5 Mean change (DDS-ADS) in number of words, number of utterances and length of utterances for dog owners, and non-owners. The positive values in the graph illustrate that on average dog-owners and non-owners produced more utterances in DDS than ADS and the negative values illustrate that on average both dog-owners and non-owners produced less words and shorter utterances in DDS than ADS. Error bars represent standard error of the mean. GLMMs revealed no effect of dog-ownership for any of the three measures.

Descriptive statistics are presented for imperatives, questions, exact repetitions, attention getting devices and greetings (Table 2.3).

Table 2.3. Mean and standard deviation for change in measures of content (DDS-ADS) for males and females (left), and dog owners and non-owners (right).

Mean (SD)	Female	Male	Dog Owner	Non-Owner
Imperatives	1.27 (2.46)	0.21 (0.43)	0.63 (1.20)	0.86 (2.47)
Questions	2.13 (1.60)	2.79 (2.55)	2.56 (2.13)	2.14 (2.14)
Exact Phrase Repetitions	3.6 (3.04)	3.43 (5.76)	4.31 (4.84)	2.36 (3.95)
Attention Getting Devices	2.13 (2.64)	2.29 (3.29)	3.06 (3.38)	1.07 (1.86)
Greetings	3.47(1.60)	4.07 (3.75)	4.31 (3.28)	2.06 (2.02)

Discussion

The results of this study showed that both males and females produced acoustically different speech in the presence of an unfamiliar dog compared with speech produced in the presence of an adult. This was demonstrated by the majority of participants producing speech with higher pitch, and greater pitch modulation when addressing the dog compared to adult-directed interactions. It did appear that females were more successful than males in producing DDS that met the criteria for pitch and pitch modulation, with 13/15 participants showing this pattern. For males, only around half of the participants produced DDS that was elevated both in pitch and pitch modulation, with most meeting just the criteria for pitch and one participant failing to meet criteria for either measures. Analysis of the change in pitch revealed a main effect of gender, with females increasing their pitch in DDS compared to ADS significantly more than males. While it also appeared that females increased their pitch modulation in DDS compared with ADS more than males, the analyses revealed that this was non-significant. In terms of dog-experience, there was no main effect and no interaction with gender. This was also the case for empathy, which did not significantly influence either pitch, or pitch modulation, and did not interact with gender. Observation of the content of DDS and ADS suggested that DDS contains fewer words than ADS and a greater number of short utterances, with increased repetition, agreeing with previous findings (Mitchell and Edmonson; 1999; Mitchel, 2001). However individual differences in gender, dog experience, or empathy had no influence on the content of speech produced. Finally, the lack of hyperarticulation of vowel sounds found in this study supports previous findings (Burnham et al, 2002; Xu et al, 2015) and is consistent with the hypothesis that hyperarticulation of vowels found in IDS facilitates language acquisition, which is not required when talking to dogs.

This study aimed to explore the individual differences that could explain acoustic variation in the production of DDS. Previous research has suggested that gender is likely to contribute to these differences, with females perhaps evolutionarily pre-disposed to producing IDS in affiliative communications. Our study confirmed that there is a gender difference in the production of DDS, with females producing a greater change in pitch between ADS and DDS compared to males. Increased pitch in IDS and DDS has previously been suggested to be important for social bonding, while the increase in pitch modulation is important for gaining the attention of the listener (Kitamura & Burnham, 2003). It is possible that this explains why in our study there was no significant difference in pitch modulation between males and females: perhaps the evolutionary predisposition of females to use IDS or PDS for affiliative communication leads to the increase in pitch, but both males and females are able to use pitch modulation as an means of gaining a listener’s attention. It would be interesting to assess the motivations underlying interactions with infants and pets, perhaps with the use of a post-trial questionnaire. This may highlight that the gender differences observed in this study may be the result of female’s increased intrinsic motivation to interact affiliatively with the listener.

From these results it does not appear that extensive experience with dogs is necessary to produce acoustically appropriate DDS, as both dog owners and non-owners were able to produce DDS increased in pitch and pitch modulation compared with ADS, and neither the acoustic properties nor the content of speech differed between groups. However, this data cannot shed much light on whether the production of DDS is something that is innate and emerges at an early age irrespective of experience, or something that requires interactions with dogs, directly or vicariously to develop. Everyone who participated was comfortable with dogs, and in our western society it is hard to avoid observing others directing DDS to pets either in reality, or on the TV. In the future, developmental research with young children and cross-cultural research in cultures where dogs are not pets but working animals may help shed light on the origins of our ability to change our speech type when interacting with dogs. Returning to the results of this study, however, the lack of significant influence of experience and empathy on the production of DDS remains interesting as it goes some way to support the idea that production of these special speech types may be universal.

Although this study found that most people modulated some acoustic aspects of their speech when interacting with an unfamiliar dog, some individuals did this to a lesser extent or in one case hardly at all. This indicates that not everyone produces dog-directed speech in a uniform manner and as the sample was necessarily self-selecting in terms of being comfortable with dogs (criterion of participation for health and safety reasons), greater variation is likely to exist in a wider sample. As subsequent chapters illustrate, DDS is useful for gaining the attention of unfamiliar dogs and increases the dogs' later affiliative preferences, so for those in roles where positive interactions are required with unfamiliar dogs (e.g. police, vets, RSPCA officers) that ability to produce DDS could be assessed and the efficacy of training to encourage this speech in those who are less able to naturally produce it could be considered.

A particularly interesting question that is outstanding from this study is whether other personality traits are likely to influence the amount or type of speech produced in this context. It is possible that the presence of an experimenter, and audio recording equipment could make the participants feel self-conscious or embarrassed, and therefore less likely to produce the type of speech they may have produced in a more natural environment. Although this study endeavoured to reduce the impact of this as much as possible by setting up the recording equipment before the participant entered the room, and the experimenter wearing headphones during the trial, it would be interesting to know whether participants who score more highly in extraversion for example would be more likely to produce DDS higher in pitch and pitch modulation compared to introverts. Other variables of interest could also include creativity and mood, as these could be assessed using pre-trial questionnaires and may highlight individual differences in speech production that have not yet been investigated.

In conclusion, this study found that males and females were able to produce DDS with elevated pitch and pitch modulation in the presence of an unfamiliar dog. A significant main effect of gender on pitch increase suggests that females alter their pitch to a greater extent than males, and this may reflect increased motivation for affiliative bonding in females, rooted in an evolutionary pre-disposition to use speech for social communication. There was no gender difference in pitch modulation, which suggests that both males and females use greater pitch modulation in DDS to increase an unfamiliar dog's attention in communicative interactions. There was no influence of experience or empathy on either pitch or pitch modulation, which suggests that these variables may not be important for determining the type of communication individuals participate in with unfamiliar dogs. Finally, I found no evidence for hyperarticulation of vowel sounds produced during interactions with an unfamiliar dog. This supports previous research that suggests hyperarticulation in IDS is important for language acquisition, and this does not occur in DDS as dogs are unable to acquire human speech. This once again supports the claim that humans are sensitive to the linguistic and affective needs of a listener, and that this is reflected in the acoustic properties of speech production.

Chapter 3: Dogs prefer naturalistic dog-directed speech

Abstract

Infant-directed speech (IDS) is a special speech type thought to aid language acquisition and improve affiliation in human infants. Although IDS shares some of its properties with dog-directed speech (DDS), it is unclear whether the production of DDS is functional, or simply an overgeneralisation of IDS within Western cultures. One recent study found that, while puppies attended more to a script read with DDS compared with adult-directed speech (ADS), adult dogs displayed no preference. In contrast, using naturalistic speech and a more ecologically valid set up, the first experiment in this study found that adult dogs attended to and showed more affiliative behaviour towards a speaker of DDS than of ADS. To explore whether this preference for DDS was modulated by the dog-specific words typically used in DDS, the acoustic features (prosody) of DDS, or a combination of the two, a second experiment was conducted. Here the stimuli from experiment 1 were produced with reversed prosody, meaning the prosody and content of ADS and DDS were mismatched. The results revealed no significant effect of speech type, or content, suggesting that it is maybe the combination of the acoustic properties *and* the dog-related content of DDS, that modulates the preference shown for naturalistic DDS. Overall, the results of this study suggest that naturalistic DDS, comprising of both dog-directed prosody, and dog-relevant content words, improves dogs' attention and may strengthen the affiliative bond between humans and their pets.

Introduction

As discussed in Chapter 1, when talking to an infant, adults use a special speech type characterised by elevated pitch, increased pitch modulation, and high affect (Burnham, Kitamura & Vollmer-Conna, 2002). This phenomenon is evident across languages including English, Russian, Swedish and Japanese (Kuhl et al, 1997; Andruski, Kuhl & Hayashi, 1999) and is thought that Infant-Directed Speech (IDS) aids language acquisition (Kuhl et al, 1997), allows selection of appropriate social partners (Schachner et al, 2011), and increases social bonding with a caregiver (Kaplan et al, 1995). As shown in Chapter 2, humans also produce a special speech type when talking to dogs (DDS) which is higher in pitch and pitch modulation than ADS, but does not contain hyperarticulation of vowel sounds which is seen in IDS (Burnham et al, 2002).

It is evident that speakers are sensitive to their audience in terms of acoustic preference, emotional needs and linguistic potential however, in order to understand the function of special speech types, it is crucial to understand how they affect the receiver. As discussed in Chapter 1, human infants show a preference for IDS from a very early age (Kaplan et al, 1995), with Cooper and Aslin (1990) finding preferences for IDS over ADS in 2 day old infants. In contrast to IDS, there has been very little research into the effect of DDS on receivers, meaning that it is currently unclear whether DDS is a non-functional overgeneralisation of IDS in Western cultures where pets often have the status of infants, or whether it functions to gain pets' attention and strengthen the affiliative bond between humans and their pets

As mentioned in Chapter 1, Ben-Aderet et al., (2017) were the first to investigate both the production of DDS, and the behavioural response to DDS in puppies, adult dogs, and older dogs. Acoustic analysis of DDS confirmed previous descriptions of the acoustic structure of this speech, where DDS was higher in pitch, with more pitch variation over time, and higher harmonicity than ADS. They also showed that human adults produced DDS to dogs of all ages. Crucially, Ben-Aderet et al., (2017) then conducted playback experiments using the DDS and ADS recorded in the first part of the study to test dog responses to these types of speech. Stimuli consisted of repetitions of the phrase *'Hi! Hello cutie! Who's a good boy? Come here! Good Boy! Yes! Come here sweetie pie! What a good boy!'* in dog- and adult-directed prosody. Speech was played from a loudspeaker in the corner of the room, with no human near the source of the sound and various measures of dogs' attention to and approach of the loudspeaker were combined into a composite behavioural response measure. They found that puppies showed a higher behavioural response to DDS than for ADS, but this preference decreased as a function of age. The authors argue that DDS may have a functional value in puppies, but not adult dogs. It is, however, possible that alternative explanations of the null result with adult dogs exist. As Ben-Aderet et al. discuss, adult dogs may need additional cues (e.g. gestures) to respond to unfamiliar speakers. If DDS functions to facilitate social communication and interaction, it may only be relevant to attend to it when it comes from a human that can be attended to and socialised with. It is possible that in this study, adult dogs realised that there is no social benefit to reacting preferentially to the speech, whereas puppies responded to DDS in the absence of a feasible producer. While it is clear that puppies are more reactive to the prosody of DDS than adult dogs, further testing with a human speaker present during stimulus presentation is required in order to rigorously test whether adult dogs really are insensitive to DDS. This study therefore aimed to test the possible function of DDS with adult dogs in a more ecologically valid setting where attention and affiliation towards the individuals who produced DDS could be directly measured. Dogs were presented with two experimenters with audio speakers on their laps that played naturalistic DDS or ADS (differing in both prosody and content) and I measured the dogs' attention to each individual during speech, and then measured the proximity to the experimenters once dogs were given the opportunity to approach them after the speech finished. I predicted that if DDS is functional for adult dogs, in experiment 1 they should attend more to DDS than ADS, and when given the opportunity to approach the experimenters they should choose to spend more time in proximity to the individual who produced DDS. I then ran a second experiment to investigate whether content or prosody was driving any preferences for naturalistic DDS. Here I presented content-mismatched stimuli (e.g. adult content with dog prosody and vice versa) and predicted that if the content of naturalistic DDS was driving preferences, dogs should attend and spend more time near the individual producing dog-relevant content. If, on the other hand, the prosody of DDS was driving preferences, as was the case for the puppies studied by Ben-Aderet et al (2017), dogs should attend and spend more time near the individual producing dog-directed prosody. Finally, if preferences for naturalistic DDS are driven by both content and prosody, or result from the combination of dog-relevant content and DDS prosody, I would expect to find no significant preference for the mismatched stimuli.

Experiment 1

As I was interested in naturalistic dog- and adult- directed speech, the stimuli used in this experiment varied in both content and prosody. The stimuli were ‘matched’ in prosody and content, such that DDS consisted of dog-relevant content and dog-directed prosody, and ADS consisted of adult-relevant content and adult-directed prosody.

Methods

Study site and participants

Dogs were recruited from Redhouse Boarding Kennels, York, with permission from the kennel owner. In experiment 1, 37 dogs took part (17 female, 20 male; mean age 6 years ± 3.86) in this study. Where dogs have been removed from various parts of the analysis due to interruptions, equipment failures, or safety reasons, the details and N for each analysis are given.

Stimuli

Stimuli were recorded as uncompressed WAV files using a Marantz PMD661 solid-state sampling at 96Hz recorder from the two human female experimenters (aged 20-21). The recordings from experimenter A were always presented through experimenter A’s speaker (and the same for experimenter B), ensuring congruency of speech with physical characteristics. Although only presenting speech from the experimenters meant that multiple dogs heard the same recordings, it ensured that the stimuli were congruous with the physical characteristics of the experimenters (age, gender, height), thus maximising ecological validity and removing the possibility of looking time measures being affected by incongruity of the stimuli. DDS was chosen from a sample of recorded naturalistic interactions with a friendly dog (irish setter).

ADS was chosen from a sample of naturalistic adult-adult interactions that occurred between the experimenters (see appendix 2 for transcripts).

Two different segments of DDS and ADS for each experimenter were selected from the continuous speech recordings (one 10 second segment and one 15 second segment). The amplitude of the speech in each segment was modified using Raven Pro (version 1.4), so that the mean RMS amplitude of each segment was equalised at approximately 3000. For each trial, the DDS track of one experimenter was paired with the ADS track of another. Figure 3.1 illustrates the stimulus timeline.

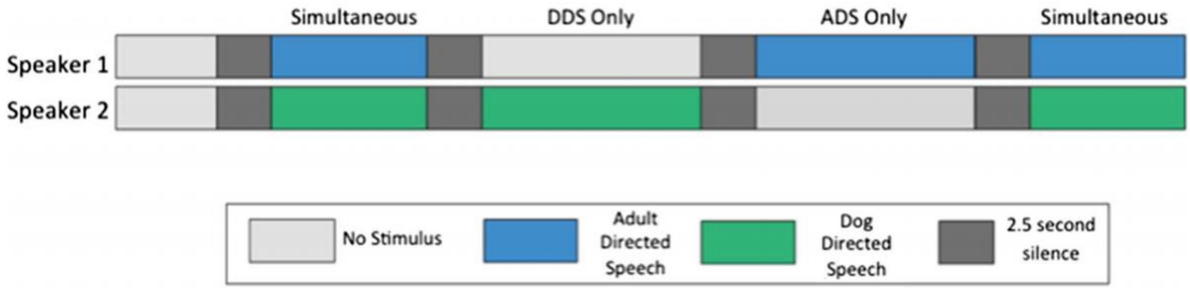


Figure 3.1. A diagram illustrating the stimulus timeline. ADS only and DDS only segments were counterbalanced such that half the dogs heard ADS only first and half heard DDS only first. Each track was played simultaneously (DDS from one speaker, ADS from another speaker) from an iPod paired with an Anchor speaker. The same 10-s segment was used in simultaneous 1 and 2 for each speaker, though these segments differed from the 15-s segments in ADS and DDS only phases.

Design

This experiment used a within subject design, where all dogs heard both DDS and ADS. All dogs heard simultaneous speech first, followed by DDS only and ADS only. The order of DDS and ADS only segments was counterbalanced across trials. Simultaneous was played again at the end, to eliminate the possibility that dogs would approach the individual who spoke last. I also counterbalanced the identity of the DDS speaker (experimenter 1 or 2), and the location from which DDS was played (left/right) across trials.

Procedure

Equipment was set up as illustrated in Figure 3.2. The speakers were equalized to 70dB at 1m away with white noise using a sound pressure meter, to ensure that that speech broadcast from each speaker would be equal in volume. Experimenters 1 and 2 then left the room via door 2. The third experimenter (handler) retrieved the dog from its kennel and entered the experimental room through door 1. The dog was allowed to explore the experimental room for 1 minute (to habituate to the environment in order to reduce distraction during the trial), before being put back on a lead and taken into a waiting room via door 3. Experimenters 1 and 2 entered through door 2 and sat in the chairs. The handler entered with the dog. Once the dog was in position the stimulus was played.

For the duration of the stimulus, the experimenters sat still to ensure the dogs were not exposed to any body language cues. The experimenters did not attempt to move their mouths simulating the speech. Instead, the experimenters placed one hand covering their mouths so that the dog could not see their lips. They also maintained neutral expressions with eyes directed towards the dog to ensure the dog did not receive differential facial cues from the experimenters.

While the stimulus played, the dog was kept on a short lead to ensure it remained within camera visibility, while still allowing the dog to move around within 1m of the handler. The handler did not interact with the dog and looked at the ground throughout. At the end of the stimulus phase, the lead was removed and the dog was allowed to explore freely for one minute and approach experimenter 1 and 2 if they wished. The dog received no interaction from any experimenter.

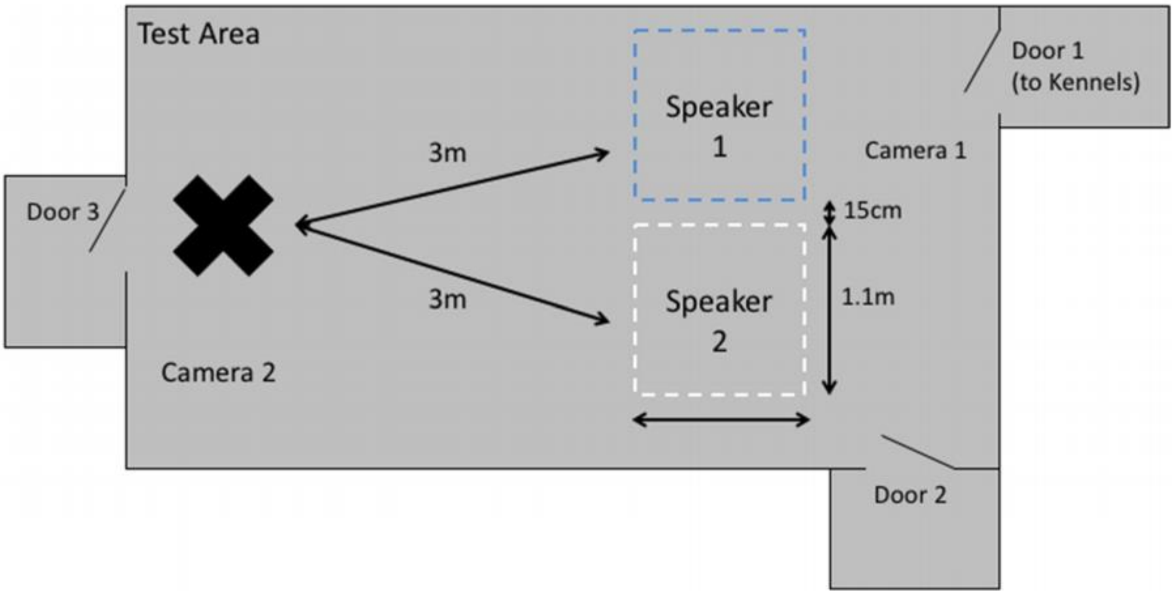


Figure 3.2. Diagram of experimental set-up at Redhouse Boarding Kennels in York. Position of dog marked with a cross. Cameras were positioned behind and to the right of the dog, and behind the speakers. Doors to other areas are marked. Dotted lines represent edges of areas in which proximity to speaker was recorded. Experimenters with speakers on their laps were seated on chairs in the centre of each area.

Video coding

Video recordings of each session were analysed and during the stimulus presentation, time spent looking towards DDS and ADS was recorded as measured by head direction. During the 1-minute off-lead period following the stimulus presentation, time spent in proximity to DDS and ADS speakers was recorded, as measured by the position of the dog's head in the 1.1m^2 area surrounding the speaker (See figure 3.2).

The period after the dog entered the room, but before the stimulus began was used as a control period (mean duration 4.56 ± 2.14 seconds). Looking times during this phase were recorded in order to establish whether the dog displayed any preference for one experimenter in particular, or one location (left or right) that may have influenced looking times in the experiment.

Inter Observer Reliability

The primary observer (AB) coded 100% of videos. For experiment 1, two trained observers each coded 30% of videos (N = 24/36 trials total) and measured looking time at each speaker in each section of the stimulus (control silence, simultaneous 1, DDS only, ADS only, simultaneous 2; N = 10 measurements) and time in proximity to each speaker in the minute post-stimulus presentation (N = 2 measurements). The primary coder had high agreement with the two secondary coders, and there was also high agreement between the two secondary coders across all measurements (Spearman's R > 0.90, *p* < .001 for all comparisons), indicating the videos had been coded reliably.

A third observer, blind to the hypotheses of the experiment also coded 22% of the videos (N = 8/36 trials total) with the sound turned off so that they were unaware which speech type was heard by the dog. There was high agreement with the primary coder for looking time (R = 0.86, *p* < .001) and for proximity preference (R = 0.96, *p* < .001).

Statistical analysis

All data were analysed using IBM SPSS (version 24) with the significance level set at *p*<.050. Attentive and affiliative preference was evaluated using Mixed ANOVAs with the fixed within-subject factor speech prosody (DDS/ADS), and between-subject factors DDS identity (Experimenter 1/ Experimenter 2), and DDS location (right/left). A single mixed ANOVA was conducted on the proximity to speakers in the minute post-stimulus presentation. For looking time, after the ANOVA on the total looking time had been completed (Table 3.1), separate ANOVAs were then run for each section of the stimulus (simultaneous; ASD only; DDS only). I applied a more conservative Bonferroni corrected alpha level to the separate section analyses (*p* = .013) to correct for family-wise error that might have arisen from running multiple tests on the same data set. Finally, an ANOVA was conducted with between-subject factors DDS identity (Experimenter 1/ Experimenter 2) and DDS location (right/left) on proportion of looking times in the control period. All assumptions of these parametric tests were tested and met.

Results

Looking Preference

For this analysis, four subjects were removed due to equipment failure (N = 33). During control silence, there was no significant main effect of Identity or Location, indicating that dogs did not display any preference for one particular experimenter or speaker location (Table 3.1).

*Table 3.1. Results of a between-subject ANOVA (df = 1,29) on looking proportions in the control period and a mixed ANOVA (df=1,29) comparing main effects and interactions for looking times towards content-matched DDS and ADS. Significant results are marked in bold where *** denotes p<.005.*

	Within Subject Effects F(*p*)				**Between Subject Effects F(*p*)**		
	Speech Type	Speech Type *Identity	Speech Type * Location	Speech Type * Identity * Location	Identity	Location	Identity * Location
Control Silence					0.38 *(.543)*	0.59 (*.448*)	0.85 (*.364)*
Total Looking	**40.51 *(<.001)**** **	0.15 *(.704)*	1.61 *(.215)*	0.24 *(.627)*	0.20 *(.656)*	1.37 *(.251)*	0.43 *(.517)*

Dogs displayed a significant preference for DDS across the whole trial (Figure 3.3; Table 3.1), and during each phase that contained DDS (Figure 3.3; Table 3.2). Dogs tended to look more towards the active ADS speaker compared to the silent DDS speaker when this was the only stimulus available, however this preference was non-significant (Figure 3.3). No significant interactions with speaker identity or location were found for total time (Table 3.1) or separate segments of the stimuli (Simultaneous, DDS only, ADS only) (Table 3.2).

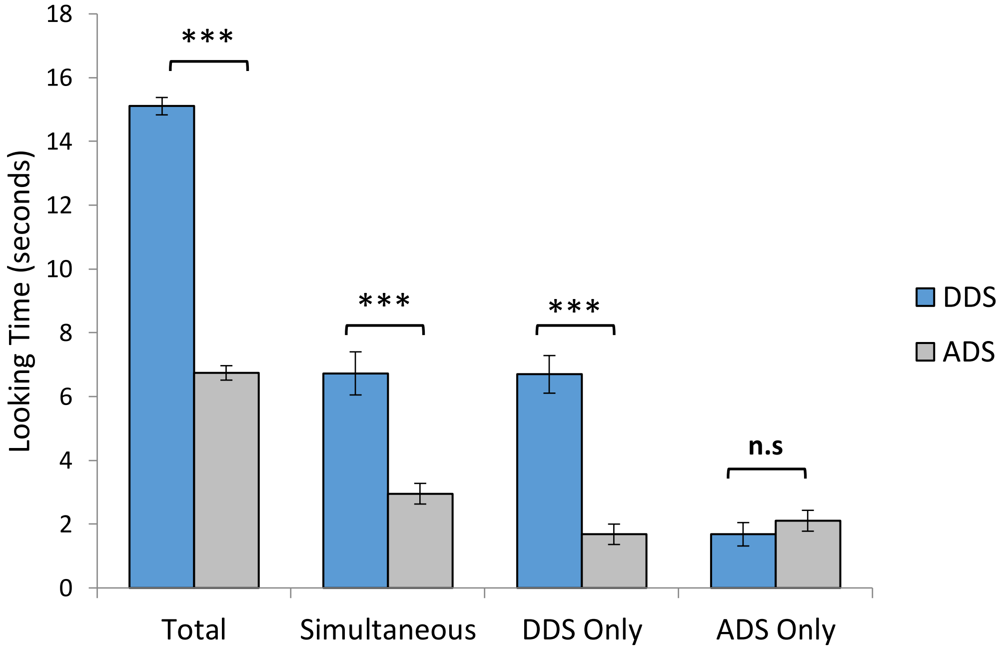


*Figure 3.3. Time spent looking towards content-matched DDS and ADS, where error bars represent 1 standard error of the mean. * denotes p<.050, ** denotes p<.010, *** denotes p<.005, and n.s represents a non-significant comparison as revealed by mixed ANOVAs.*

*Table 3.2. Results of a between subjects ANOVA (1,29) for looking time during the control period and a Mixed ANOVA with Degrees of Freedom (1,29) comparing main effects and interactions for looking times towards content-matched DDS and ADS. Significant results are marked, where * = p<.050 and ** = p< .005 for total looking time and control silence. Bonferroni corrected alpha was employed for individual segments of the stimulus (p=.013). No significant effects or interactions were found at this level.*

	Within Subject Effects F(*p*)				**Between Subject Effects F(*p*)**		
	Speech Type	Speech Type *Identity	Speech Type * Location	Speech Type * Identity * Location	Identity	Location	Identity * Location
Control Silence					0.38 *(.543)*	0.59 (*.448*)	0.85 (*.364)*
Simultaneous	**30.12 *(<.001)*****	1.43 *(.241)*	2.76 *(.107)*	<0.01 *(959)*	0.70 *(.408)*	0.08 *(.785)*	0.17 *(.685)*
DDS Only	**48.46 *(<.001)*****	0.04 *(.841)*	0.35 *(.562)*	0.00 *(.961)*	2.83 *(.103)*	4.54 *(.042)*	0.36 *(.551)*
ADS Only	3.81 *(.061)*	*2.29 (.141)*	1.92 *(.176)*	0.03 *(.855)*	0.37 *(.548)*	0.16 *(.692)*	0.26 *(.614)*
Total Looking	**40.51 *(<.001)******	0.15 *(.704)*	1.61 *(.215)*	0.24 *(.627)*	0.20 *(.656)*	1.37 *(.251)*	0.43 *(.517)*

Proximity Preference

For this analysis, three dogs were removed from the data set due to equipment failure or because the dog had to be kept on a lead, resulting in an N = 34. A repeated measures ANOVA revealed that after hearing content matched stimuli, dogs spent significantly more time in close proximity to the DDS speaker than the ADS speaker (F (1, 30) = 5.54, *p* = .025*;* Figure 3.4). No significant interactions with location or speaker identity were found (Table 3.3).

Table 3.3. Results of a mixed ANOVA with degrees of freedom (1, 30) comparing the time spent near DDS and ADS speakers for content-matched speech.

	Within Subject Effects F(*p*)				**Between Subject Effects F(*p*)**		
	Speech Type	Speech Type *Identity	Speech Type * Location	Speech Type * Identity * Location	Identity	Location	Identity * Location
Proximity Preference	**5.54(*.025*)***	1.64 (*.210*)	0.29 (*.592*)	0.05 (*.833*)	1.13 (*.552*)	0.36 (*.552)*	0.62 (*.438*)

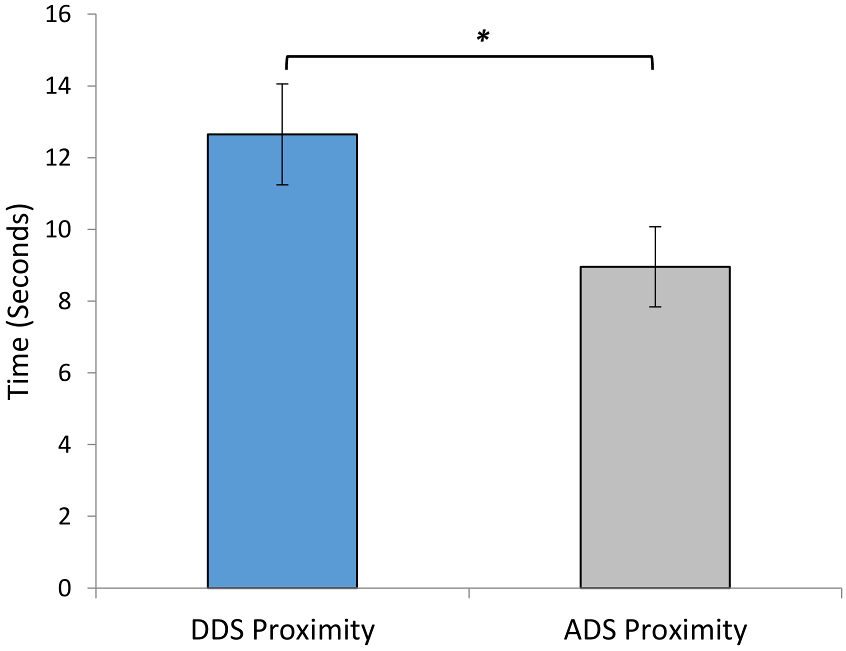


*Figure 3.4. A graph to show the mean time spent in proximity to each experimenter (seconds), in the minute after the speech stimuli ended, when the dogs heard content-matched DDS and ADS. Error bars represent one standard error of the mean. * denotes a significant main effect of speech type where p<.050 based on the results of a mixed ANOVA.*

Content-matched speech: Discussion

This experiment showed that dogs display a behavioural preference for naturalistic DDS (matched in prosody and content) compared with ADS when presented in the presence of an associated human. Dogs, on average, spent more time looking towards a speaker of DDS compared with a speaker of ADS in all segments of the stimulus containing DDS and across the trial as a whole. It was also found that when given the subsequent opportunity to interact with the speakers, dogs chose to spend more time in proximity with the DDS speaker, than the ADS speaker. Although the absolute differences in looking and proximity time were small and therefore their functional relevance may be questioned, I feel the substantial effect sizes obtained and the convergence of results across the behavioural measures indicates that this study has detected functionally relevant differences in behaviour. Overall, these results support the hypothesis that dogs display attentive and affiliative preferences for naturalistic DDS over ADS.

The results from the control period show no significant preference for a specific location, or speaker identity, indicating that the dogs had no a priori preference for looking at one experimenter or location. In line with this, no significant main effects of location or speaker identity, or interactions of identity, location, and speech type were found.

Although these results show a robust preference for naturalistic DDS over ADS, as the stimuli in this experiment differed in both content and prosody it is not possible to determine whether this effect is driven by dog-directed prosody or content, as these factors did not vary independently. Therefore, although this experiment clearly shows that dogs discriminate between, and show a behavioural preference for, naturalistic DDS over ADS, further investigation is required to determine the extent to which prosody and content are driving this preference.

Experiment 2: Content-mismatched speech

Experiment 2 was designed in order to examine whether content alone, or prosody alone was sufficient for driving the preference found in experiment 1. In experiment 2 the content from experiment 1 was reproduced but with reversed prosody such that the dog-related content was spoken with the prosody of ADS and vice versa. For simplicity, in all cases, DDS refers to stimuli with dog-directed prosody (with either dog or adult related content) and ADS refers to stimuli with adult-directed prosody (with either adult or dog related content). In Experiment 2 dogs were presented with content-mismatched DDS (dog-directed prosody with adult-related content) and content-mismatched ADS (adult-directed prosody with dog-related content).

Methods

Study site and participants

In experiment 2, 32 dogs from Redhouse Boarding Kennels in York took part (16 female, 16 male; mean age 6 years ± 3.75).

Stimuli

For experiment 2, uncompressed wav. files were recorded from two new female experimenters (age 20 and 21). The experimenters repeated the transcripts from experiment 1 with the opposing prosody, in order to produce content-mismatched DDS and ADS. All stimuli were still directed to an appropriate live audience (e.g. adult script was produced with dog prosody to a live dog; Irish setter) and processed as described in experiment 1.

For the stimuli used in experiment 2, some dog content was repeated in ADS, and some adult content was removed in DDS. This was in order to account for differences in word rate between naturalistic DDS and ADS, These alterations are indicated in Appendix 2. The amplitude of the speech segments was again equalised and tracks were built as in experiment 1 (see figure 3.1).

Acoustic Analysis of stimuli

To ensure the prosody of the content-mismatched DDS and ADS for experiment 2 were convincing, I compared the acoustic properties of these stimuli with the stimuli used in experiment 1. Mean, minimum and maximum pitch (FO) was measured (Table 3.4) in PRAAT (version 6.0.05). Pitch settings were 75-1200Hz and continuous segments of speech with a continuous visible pitch line were selected and the mean, min and max pitch in the segment was extracted using the ‘get pitch’ function. Pitch modulation was calculated as maxF0-minF0.

Word rate was calculated as the number of words divided by the duration from the start of the first word to the end of the last word in a stimulus.

Table 3.4. Acoustic measurements of the different types of speech produced by each experimenter.

Speaker ID	Prosody	Content	Mean Pitch	Pitch Modulation	Word Rate
Experimenter 1	DDS	Dog	598.88	240.26	172.85
	ADS	Adult	452.68	170.02	216.01
Experimenter 2	DDS	Dog	794.51	207.49	195.37
	ADS	Adult	413.47	62.97	242.40
Experimenter 3	DDS	Adult	684.58	285.92	138.97
	ADS	Dog	487.00	87.45	270.53
Experimenter 4	DDS	Adult	535.02	172.18	128.95
	ADS	Dog	472.75	83.26	278.71

Generalized Linear Mixed Models (GLMMs) were conducted to assess the effect of prosody (dog-directed/adult-directed); content (dog/adult) and content-prosody matching (matched (experiment 1)/mismatched (experiment 2)) on the acoustic measurements of stimuli in experiments 1 and 2. These factors were entered as fixed factors in models with (i) mean pitch and (ii) pitch modulation as DVs. In order to ensure I was comparing the pitch-related measures of the same words or phrases, for mean pitch and pitch modulation, measurements of each continuous segment of speech with a continuous visible pitch line that were available in both experiments were entered into the analyses. Each speech segment was numbered and included as a random factor along with speaker identity, in order to control for repeated sampling at these two levels (Waller et al, 2013). For word rate, the rate of each 10 or 15 second stimulus produced by each speaker was entered into analyses, with speaker identity entered as a random factor to control for repeated sampling of each speaker. As there were only a small number of data points for this GLMM (N = 16), I ran three separate models, each with a single fixed factor (prosody, content or prosody-content matching) to avoid overfitting the models.

GLMMs revealed that the content-matched (experiment 1) and content-mismatched stimuli (experiment 2) did not significantly differ in pitch, pitch modulation or word rate (Table 3.4; Table 3.5), indicating that the content mis-matched stimuli were produced with prosody representative of natural dog-directed and adult-directed speech. In line with previous descriptions of the prosody of DDS, the pitch was significantly higher, the pitch modulation significantly greater, and word rate significantly slower for stimuli produced with dog-directed prosody compared to adult-directed prosody (Burnham et al, 1998; Ben-Aderet et al., 2017; Table 3.4; Table 3.5). Content did not significantly affect pitch modulation or word rate, but dog-content was significantly higher pitched than adult-content (Table 3.4; 3.5).

*Table 3.5. Results of GLMMs exploring the effect of prosody, content, and content-prosody matching, on pitch, pitch modulation, and word rate. Significant results are indicated, where *** denotes p<.001.*

	df	**Prosody** F(*p*)	**Content** F(*p*)	**Content-Prosody Matching** F(*p*)
Mean Pitch	1, 328	**245.86** (<.001)***	**13.97**(<.001)***	0.58(.447)
Pitch Modulation	1, 328	**49.13** (<.001)***	0.07 (.792)	0.20 (.653)
Word Rate	1, 6	**34.22** (<001)***	3.24 (.094)	<0.01 (.937)

Design

As in experiment 1, this experiment used a within subject design with all dogs hearing both DDS and ADS. Between subject factors such as DDS speaker, DDS location, and stimulus order were counterbalanced across trials.

Procedure

The procedure for this experiment was identical to that of experiment 1.

Inter Observer Reliability

The primary observer (AB) coded 100% of videos. Two trained observers each coded 50% of the videos (N = 32/32 trials total). The primary observer had high agreement with both secondary coders, who also had high agreement with each other across all measurements (Spearman's R > 0.90, *p* < .001 for all comparisons).

A third observer, blind to the hypotheses of the experiment also coded 22% of the videos (N = 7/32 trials total) with the sound turned off so that they were unaware which speech type was heard by the dog. There was high agreement with the primary coder for looking time (R = 0.93, *p* < .001) and for proximity preference (R = 0.88, *p* < .001).

Statistical analysis

As above, attentive and affiliative preference was evaluated using Mixed ANOVAs with the fixed within-subject factor speech prosody (DDS/ADS), and between-subject factors DDS identity (e.g. Experimenter 1/ Experimenter 2), DDS location (right/left), and stimulus order (DDS 1st / ADS 1st). All assumptions were tested and met.

Experiment 2: Results

Looking Preference

For content-mismatched DDS, 3 trials were removed due to equipment failure and the following analysis is based on n = 29. A mixed ANOVA revealed there was no significant preference for DDS when content was incongruent with prosody (Figure 3.5; Table 3.6; Table 3.7). During the control period, there was a main effect of identity, with dogs preferring to look towards Experimenter 3 compared to experimenter 4 (Table 3.6).

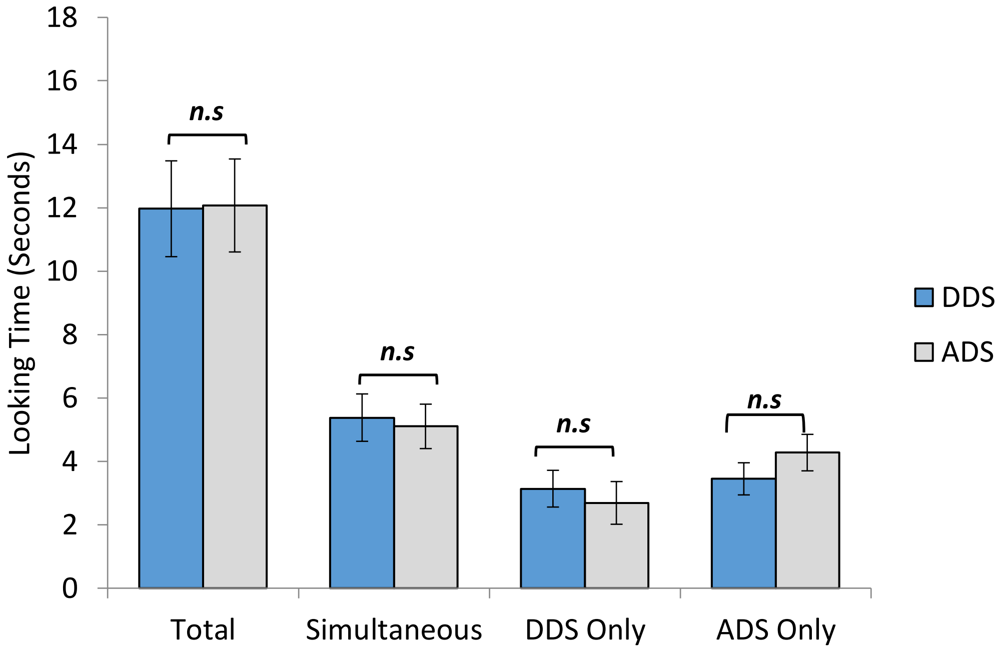


Figure 3.5. Time spent looking towards content-mismatched DDS and ADS during each phase, where error bars represent 1 standard error of the mean. N.s denotes non-significant comparisons as revealed by mixed ANOVAS.

*Table 3.6. Results of between subject ANOVA (1, 25) for the control silence and a mixed ANOVA with Degrees of Freedom (1,25) comparing main effects and interactions for looking times towards content-mismatched DDS and ADS. Significant results are marked, where * indicates p<.050.*

Within Subject Effects F(*p*)					**Between Subject Effects F(*p*)**		
	Speech Type	Speech Type *Identity	Speech Type * Location	Speech Type * Identity * Location	Identity	Location	Identity * Location
Control Silence					**4.24 (*.048*)***	1.44 (*.242*)	1.02 (*.322*)
Total Looking	<0.01 (*.985*)	**5.75 (*.024*)***	2.03 (*.167*)	1.00 (*.328*)	2.58 (*.121*)	0.99 (*.330*)	0.34 (*.560*)

There was also an interaction of speech type and identity for total looking time. To explore the nature of the interaction between speech type and identity, four independent samples t-tests with Bonferroni corrected alpha (*p* < .013*)* were conducted. Firstly, at the level of DDS, there was a significant main effect of speaker identity, with dogs preferring the speech of experimenter 3 over experimenter 4 (t (27) = 3.08, *p* = .005). However, at the level of ADS, there was no significant effect of speaker identity (t (27) = 0.82, *p* = .419. At the level of each speaker, there was no preference for the DDS of experimenter 3 compared with her ADS (t (27) = 0.77, *p* = .450), and the same was true for experimenter 4 (t (27) = -1.50, *p* = .146).

*Table 3.7. Results of Mixed ANOVA with Degrees of Freedom (1,25) comparing main effects and interactions for looking times towards content-mismatched DDS and ADS. Significant results are marked, where * = p<.050 and ** = p< .005 for control silence and total looking time. Bonferroni corrected alpha was employed for individual segments of the stimulus (p=.0125). No significant effects or interactions were found at this level.*

	Within Subject Effects F(*p*)				**Between Subject Effects F(*p*)**		
	Speech Type	Speech Type *Identity	Speech Type * Location	Speech Type * Identity * Location	Identity	Location	Identity * Location
Control Silence					**4.24 (*.048*)***	1.44 (*.242)*	1.02 (*.322*)
Simultaneous	0.18 (*.678*)	2.15 (*.155*)	2.51 (*.126*)	2.53 (*.125*)	0.53 (*.474*)	0.23 (*.635*)	0.02 (*.899*)
DDS Only	0.15 (*.702*)	6.33 (*.019*)	0.10 (*.756*)	0.09 (*.765*)	1.80 (*.192*)	0.93 (*.345*)	0.24 (*.628*)
ADS Only	1.02 (*.322*)	1.77 (*.195*)	1.34 (*.257*)	0.56 (*.463*)	3.38 (*.078*)	0.90 (*.351*)	1.30 (*.266*)
Total Looking	<0.01 (*.985*)	**5.75 (*.024*)***	2.03 (*.167*)	1.00 (*.328*)	2.58 (*.121*)	0.99 (*.330*)	0.34 (*.560*)

Proximity Preference

This analysis is based on N=30 following equipment failures. For content-mismatched stimuli, dogs spent more time, on average, in proximity to the ADS location as illustrated by figure 3.6. However a mixed ANOVA revealed that this result was non-significant (See Table 3.8).

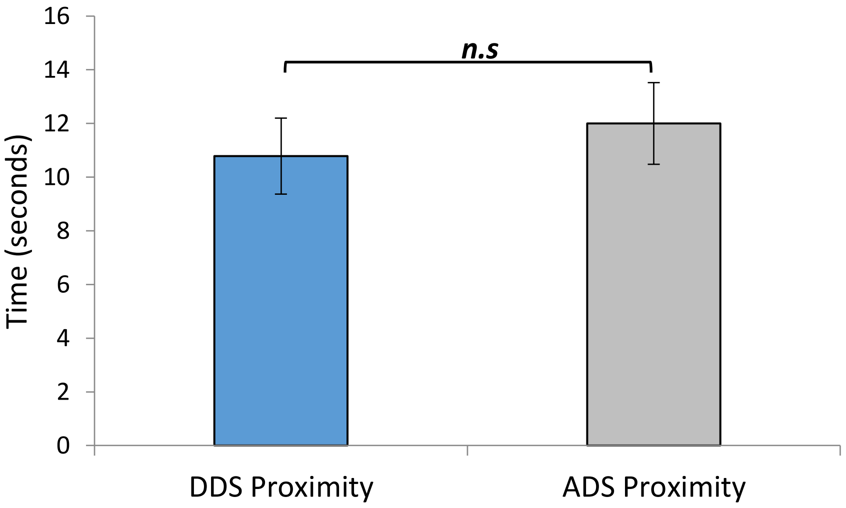


Figure 3.6. A graph to show mean time spent in proximity with each speaker (seconds), for content-mismatched DDS and ADS. Error bars represent one standard error of the mean.

Table 3.8. Results of a Repeated Measures ANOVA with degrees of freedom (1, 26) comparing the time spent near DDS and ADS speakers for content-mismatched speech.

	Within Subject Effects F(*p*)				**Between Subject Effects F(*p*)**		
	Speech Type	Speech Type *Identity	Speech Type * Location	Speech Type * Identity * Location	Identity	Location	Identity * Location
Proximity Preference	1.03 (*.319*)	0.85 (*.365*)	0.01 (*.992*)	0.02 (*.894*)	1.20 (*.283*)	0.59 (*.448)*	0.52 (*.477*)

To explore whether the failure to find a significant preference for either type of speech was likely due to reduced power associated with the slightly smaller sample size in experiment 2 compared to experiment 1, I considered effect sizes and conducted power analyses using G*Power (V3.1.9.2). The preference for attending to DDS in experiment 1 was associated with a large effect size ($\eta^2 = 0.563$), yet the same comparison in experiment 2 yielded a very small effect size ($\eta^2 < 0.001$). An a-priori power analysis for looking time in experiment 2 indicated that to find a similar effect size based on partial η^2 of 0.56, with power of 0.80 and an alpha level of 0.05 for the within-subjects comparison of speech type, 6 participants would have been needed, which was exceeded with 29 participants in experiment 2. The proximity preference for the DDS speaker in experiment 1 was associated with a medium effect size ($\eta^2 < 0.156$), yet the same comparison in experiment 2 yielded a small effect size ($\eta^2 = 0.038$). An a-priori power analysis for proximity duration in experiment 2 indicated that to find a similar effect size based on partial η^2 of 0.16, with power of 0.80 and an alpha level of 0.05 for the within-subjects comparison of speech type, 24 participants would have been needed, which was exceeded with 30 participants in experiment 2. Together the effect sizes and power analysis indicates that experiment 2 had sufficient power to find differences similar to those found in experiment 1, had they existed, and therefore I can be relatively confident in this null result.

Content-mismatched speech: Discussion

The results from experiment 2 suggest that there is no significant difference in dogs' attention or proximity preference to speakers of DDS or ADS where content and prosody did not match. This suggests that neither content, nor prosody, is solely responsible for the preference for DDS shown in experiment 1. As the same scripts were used in both experiments, this result also highlights that the preference shown in experiment 1 could not be explained by the use of specific words in the content of the original stimuli, such as 'walk' or 'dog', for example. If this were the case, it would have been possible to observe a preference for content-mismatched ADS, which not only contained the specific dog-related words used in experiment 1, but more repetitions of them (see methods).

In order to explore alternative explanations for these null results I first considered if the difficulty of producing these content mismatched stimuli had resulted in poor examples of DDS and ADS prosody being produced. The acoustic analysis of the stimuli, however, illustrates that the content-mismatched stimuli followed the same patterns of acoustic properties as the naturalistic DDS of experiment 1. This supports the use of these stimuli, and highlights that the null result found in this experiment is unlikely to be due to failures in producing authentic DDS or ADS when the content is reversed. Second, although a broadly comparable number of subjects were used in experiments 1 and 2, it is possible that the slightly smaller N available in Experiment 2 (33 vs 29 Looking duration; 34 vs 30 proximity duration), left experiment 2 with slightly less power to detect differences compared to Experiment 1. However, examination of effect sizes indicates that whilst the naturalistic speech in experiment 1 elicited large effect size ($\eta^2 = 0.563$), effect sizes obtained with the reversed stimuli were extremely small ($\eta^2 < 0.001$). Power analyses confirmed that this study had sufficient sample sizes in experiment 2 to detect differences similar to those found in experiment 1. I am therefore confident that the null result in experiment 2 was not due to lack of power.

In experiment 2 a significant interaction between speech type and experimenter revealed that Experimenter 3's DDS was more effective at eliciting attention than Experimenter 4's DDS. This effect is likely mediated by what seemed to be an *a priori* preference for Experimenter 1, which resulted in dogs looking significantly longer at this experimenter in the control period before any speech was produced. It is not clear whether visual or scent characteristics drove this preference, although scent seems unlikely as the preference did not remain in the post-stimulus proximity to experimenters where an attractive scent could have been actively explored. It is interesting that dogs seemed to have an immediate preference for one experimenter and this may have enhanced the efficacy of an experimenter's dog-directed prosody. It is, however, important to note that the preferred experimenter's DDS was still not significantly more effective in attracting dogs' attention than her ADS. Indeed post-hoc analyses of the interaction term at the level of each speaker confirmed the main findings that the different types of speech did not elicit significantly different behaviour from the dogs.

Discussion

The results provide evidence that in an ecologically valid setting, dogs attended more towards naturalistic DDS, where prosody and content were matched, compared with ADS. This study also shows for the first time that dogs subsequently spend more time in proximity to an experimenter who has recently produced naturalistic DDS than one who has recently produced ADS. This novel finding suggests that DDS may fulfil a dual function of improving attention and increasing social bonding. This fits with the current understanding of infant research, which suggests that IDS not only serves to facilitate language acquisition, but that it is also crucial for developing meaningful social relationships with caregivers.

The second experiment was designed to investigate whether prosody or content alone was driving this preference for naturalistic DDS. However, when content and prosody were mismatched, there was no difference in the amount of time spent in proximity to the experimenters and there was no significant attentive preference for DDS or ADS in any part of the trial, or across the session as a whole. This suggests that neither content nor prosody alone was driving the preference observed in experiment 1. Instead, it is clear that both content and prosody matter to dogs. Further, the results from Ben-Aderet et al. (2017), demonstrated that adult dogs do not prefer dog-relevant content produced with dog-directed prosody over adult-directed prosody, confirming that it is e the congruent combination of dog-directed content and prosody that underpins the preference for naturalistic DDS.

Interestingly, Ben-Aderet et al. did find a significant preference for DDS prosody in puppies, showing that puppies are more sensitive to prosodic differences compared to adult dogs. Puppies may be more sensitive to acoustic differences than adult dogs in the same way that human babies are most sensitive to IDS early in life (Newman & Hussain, 2006). Puppies also have less experience of human language and time to form associations between specific words and positive experience (e.g. walk), and thus should be less sensitive to content. Therefore, whilst puppies may rely wholly on prosodic information, adult dogs seem to take both content and prosody into account, and only when these two things are relevant to them do they display a behavioural preference. Whilst preference for dog-related content needs experience of human-interaction to develop, the origins of the preference for dog-directed prosody are less clear: they may be routed in an innate preference for higher pitched, tonal sounds or be a product of early learning environment. If preferences for DDS prosody are based on preferences for high pitched tonal sound, which across mammalian species is associated with affiliation and submission rather than aggression (Morton, 1977), then other mammalian species should show a preference for DDS over ADS. Future research could test this possibility. Alternatively, preference for DDS prosody may have arisen through various routes during the domestication process. Firstly, early in the domestication process, DDS may have provided dogs with a reliable cue that indicates safe social partners at a time when joining human groups may have been dangerous, and identifying those who would not be hostile would have been important for a dog's survival. Secondly, as dogs are able to engage with humans in joint attention (Miklósi et al., 2003), and can cooperate to achieve goal-directed actions (Range & Virányi, 2014), it is possible that humans selected dogs for characteristics that promoted social communication during domestication, including attentive and affiliative preference for DDS. It is, however, also possible that dogs kept as pets are conditioned over their individual lifetimes to respond positively to DDS as this type of speech is often paired with positive events (e.g. food treat, toy, walk or affection). Although Ben-Aderet et al. found a clear preference for DDS in young dogs (2-5 months), it is possible that such associations could be formed in that time. Future research with young puppies raised with extremely minimal human contact, would enable us to test whether environmental input is needed to shape this preference or whether it is an innate preference, as it seems to be in human infants (Cooper & Aslin, 1990).

Although the use of real people to deliver the speech to the dogs increased the ecological validity of the experimental set up, it did have potential drawbacks. First, the importance of providing speech from each experimenter (exact match with characteristics including gender, height, and size) to ensure it was physically congruous meant that the same stimuli were heard by multiple dogs. Although acoustic analysis confirmed the structure of these stimuli were representative of DDS and ADS reported in other studies, it is unclear whether these findings would generalise to a wider sample of DDS and the findings suggest that there may be individual variation in the efficacy of DDS. Thus, further studies without pseudoreplication at the level of the stimulus are required to confirm the generalisability of the findings. Differential a priori interest in the experimenters, as was found in Experiment 2, is a further complication associated with the use of live models in these experiments, which highlights the need for rigorous counterbalancing and a control period where such *a priori* biases can be measured. In addition, the results illustrate the interesting possibility that *a priori* preferences for individuals may influence the effectiveness of, and sensitivity to other cues including speech type.

In conclusion, the results from this study support the hypothesis that dogs pay more attention to naturalistic DDS than to ADS. It also revealed that dogs spent more time near someone who had just produced DDS rather than ADS, indicating for the first time that DDS may not just modulate attentive behaviour, but also play a role in the development of affiliative preferences. This preference for naturalistic DDS was not driven by preference for dog-directed content or prosody alone, as no attentive or affiliative preferences were shown when dogs were presented with content and prosody mismatched stimuli. This study concludes that naturalistic DDS elicits more attention from dogs than ADS and has the potential to strengthen the affiliative bond a human has with a dog.

Chapter 4: The interaction of speech type and eye-gaze in communications with dogs

Abstract

Previous research has shown that human infants display a behavioural preference for direct eye-gaze from a very young age and that mutual gaze is an important communicative cue that a partner's behaviour should be attended to. It is also known that sharing mutual gaze is an important component of attachment, and forms the foundations of affiliative bonding in mother-infant relationships. Many animals avoid direct eye-contact with humans and conspecifics, whereas dogs look readily and frequently to human faces for social and communicative cues. This study aimed to investigate dogs' preference for direct compared with averted eye-gaze, and the influence of gazing cues on the preference for dog-directed speech. Dogs heard adult-directed speech (ADS), and dog-directed speech (DDS) from two experimenters who alternated between direct and averted eye-gaze. The results show that dogs preferred to attend to DDS over ADS, and direct over averted eye-gaze. Importantly, there was also an interaction of speech type and eye-gaze, with direct eye-gaze selectively enhancing the preference for DDS, but not ADS. These results suggest that dogs are sensitive to eye-gaze cues given by humans, and that looking at a dog while also using DDS may improve communications with unfamiliar dogs, in a similar way to human infants.

Introduction

In mother-infant relationships, the sharing of mutual gaze is thought to be a fundamental component of attachment, forming the basis of affiliative bonding, and helping to establish communicative links (Dickstein, Thompson, Estes, Malkin & Lamb, 1984; Farroni, Csibra, Simion & Johnson, 2002). From birth, human infants show a preference for faces with open eyes (Batki, Baron-Cohen, Wheelwright, Connellan & Ahluwalia, 2000), and prefer to look towards faces that engage with them in direct eye contact. From as early as 2-5 days old, infants show enhanced neural processing of direct gaze compared to averted gaze, and Farroni et al., argue that this exceptionally early sensitivity to mutual gaze is the foundation for the development of social skills in humans. Perception of averted gaze causes an automatic shift of attention in the direction of the other's focus (Driver et al., 1999) which is proposed to form the foundation of joint attention (Butterworth & Jarret, 1991). Importantly, mutual eye-gaze is thought to be an important ostensive cue for establishing communicative context between humans (Kleinke, 1986; Hains & Muir, 1996; Symons, Hains & Muir, 1998,) providing a consistent cue to infants that a communication is meant for them, and that they should therefore pay attention. This helps infants to attend to behaviours that are important for learning and development, a phenomenon known as natural pedagogy (Csibra & Gergely, 2009) which is thought to occur reliably in human infants from at least 4 months of age. It is not clear whether the interpretation of eye-gaze signals as a communicative signal is an adaptation that is unique to humans (Povinelli & Giambrone, 2001). However, in recent years evidence has emerged of other species engaging in mutual gaze with humans. As discussed in Chapter 1, Ferrari, Kohler, Fogassi and Gallese (2000) showed that adult pig-tailed macaques were able to use the gaze cues of a human experimenter to locate hidden food, while infant macaques required more information from head-direction or a combination of these cues in order to perform above chance. These results showed that non-human primates are able to use human communicative cues including gaze, but only in limited contexts, and far less reliably than human infants, at least across species.

In relation to human-given cues, there is evidence that dogs are sensitive to human gestures from a young age and as adults they can be highly proficient using human-given cues including pointing, bowing, nodding, head-turning, and glancing to help them successfully locate hidden food (Miklosi, Polgardi, Topal & Csanyi, 1998). As discussed in Chapter 1, even highly socialised wolves are unable to perform as well as dogs on these tasks (Bentosela et al, 2016), and do not look towards human faces when faced with an impossible task, where domesticated dogs do so readily (Miklosi et al, 2003). Miklosi et al., argue that through positive feedback, either evolutionarily or ontogenetically, the readiness of dogs to look at human faces may have led to the development of complex human-dog communication that cannot be achieved in wolves.

In dogs it is also possible to observe biological changes in response to sharing mutual eye-gaze with humans, reflective of changes seen in mother-infant relationships. With human infants, increased duration of mutual eye-gaze is linked to increased levels of maternal oxytocin (Kim, Fonagy, Koos, Dorsett & Strathearn, 2014). This has been linked with the oxytocin positive loop, which increases maternal nurturing, and subsequently increases oxytocin in the infant (Nagasawa, Okabe, Mogi & Kikusiu, 2012). Dogs appear to engage in this oxytocin positive loop in cross-species relationships with humans (Nagasawa et al, 2015). While mutual gaze is clearly important for affiliative bonding with infants and with dogs, it rarely occurs in isolation in natural mother-infant interactions in humans, and is instead thought to be linked to the use of verbal communication (Tronick et al, 1978). Haith, Bergman and Moor (1977) recorded three 11-week-old infants as they scanned an adult's face while he/she was stationary, moving, or talking. Talking while looking at the infant led to greater fixation on the eye area in 7-11 week-olds. This suggests that there may be some interaction between eye-gaze and speech that is important to communication and social development in human infants. In previous chapters I have discussed the importance of speech type in language learning and the building of social relationships for human infants, and in Chapter 3 I found that dog-directed speech (DDS) seems to be important in developing social bonds with humans. In my previous experiment, direct eye-gaze was always present during speech, so it was not possible to probe how eye-gaze contributed to the strong preference for DDS the dogs displayed.

This study aims to investigate whether eye-gaze enhances responsivity to DDS using an ecologically valid paradigm based on that described in Chapter 3. I have already shown that dogs display a strong behavioural preference for DDS compared to adult-directed speech (ADS) (Benjamin & Slocombe, 2018) and aim to replicate this finding here. I hypothesise that dogs will look more towards, and choose to spend more time in proximity with a speaker of DDS than one of ADS, reflected in a main effect of speech type. I also expect dogs to prefer to look towards an experimenter whose gaze is directed towards them, and so I predict a main effect of eye-gaze for looking duration. I do not expect to find a main effect of eye-gaze on proximity to speaker as all dogs will experience direct and averted eye-gaze from both speakers. Based on the evidence from human infants (Tronick et al, 1978; Haith, Bergman and Moor, 1977) I expect that mutual eye-gaze will provide dogs with a cue that the accompanying speech is meant for them and therefore direct eye-gaze should enhance the preference for DDS. I therefore hypothesise that as well as preferences for DDS, and eye-gaze separately, there will be an interaction of speech type and eye-gaze, where engaging in direct gaze will selectively enhance attention for DDS. More specifically, I predict that eye-gaze during DDS will generate an increase in attention towards the DDS speaker compared with the ADS speaker, whereas, eye-gaze during ADS will not produce a significant increase in attention towards the ADS speaker compared to the DDS speaker.

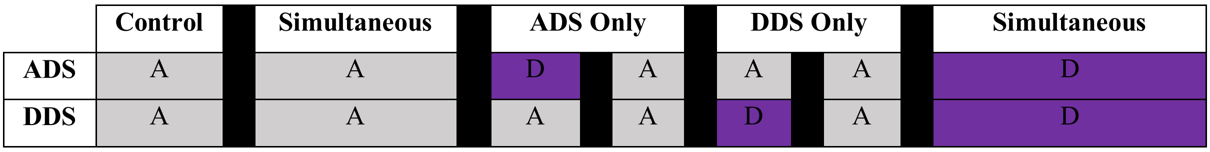
Methods

Study site and participants

Dogs were recruited from Redhouse Boarding Kennels, York, with permission from the kennel owner, and the Waggy Dog Crèche, York, with permission from the crèche and the dog's owners. In total, 52 dogs took part (16 females and 36 males; mean age 3 years ± 2.58, 29 from Waggy Dog Creche, 11 from RH) in this study. Where dogs have been removed from various parts of the analysis due to interruptions, equipment failures or safety reasons, the details and N for each analysis are given.

Stimuli

Stimuli were recorded as uncompressed WAV files using a Marantz PMD661 solid-state recorder from the two human female experimenters (aged 20–21). As described in Chapter 3, the recordings from experimenter A were always presented through experimenter A's speaker (and the same for experimenter B), ensuring congruency of speech with physical characteristics. Although this led to multiple dogs hearing the same recordings, it ensured that the stimuli were congruous with the physical characteristics of the experimenters (age, gender, height), thus maximising ecological validity and removing the possibility of looking time measures being influenced by incongruity of the stimuli. To control for content words, experimenters read scripted sentences of dog- and adult-directed speech. DDS stimuli were recorded when the scripts were addressed to a friendly dog (Parson-Russell Terrier) and ADS stimuli were recorded when the scripts were addressed to another experimenter. Three different segments of DDS and ADS for each experimenter were produced, each 10 seconds in length. The amplitude of the speech in each segment was modified using Raven Pro (version 1.4), so that the mean RMS amplitude of each segment was equalised to 3013.17 ± 172.27. For each trial, the DDS track of one experimenter was paired with the ADS track of another. Figure 4.1 illustrates the stimulus timeline.



	Control	■	Simultaneous	■	ADS Only			■	DDS Only			■	Simultaneous
ADS	A	■	A	■	D	■	A	■	A	■	A	■	D
DDS	A	■	A	■	A	■	A	■	D	■	A	■	D

Figure 4.1. Stimulus track timeline. Each segment of speech is 10 seconds in length, with 2.5 seconds of silence between (black), (control silence mean duration 6.41 ± 2.45 seconds). During simultaneous segments, both experimenters displayed the same eye-gaze, either both direct (D, purple), or both averted (A, grey). Both experimenters looked down during silence between segments. Order of ADS only, and DDS only segments were counterbalanced across dogs, as well as eye-gaze order, resulting in 32 stimulus sets

Design

This experiment used a within-subject design, where all dogs experienced all combinations of speech type (DDS/ADS), and eye-gaze (Direct/Averted). All dogs heard simultaneous speech first, followed by DDS only and ADS only. The order of DDS and ADS only segments was counterbalanced across trials. I also counterbalanced the order of eye-gaze (within segments), although the same order was maintained in DDS only and ADS only segments such that both speakers performed direct then averted gaze, or both performed averted then direct. Gaze during simultaneous segments was always matched between speakers (both direct, or both averted). These counterbalancing procedures led to 32 combinations of gaze and speech type. Simultaneous was played again at the end, to eliminate the possibility that dogs would approach the individual who spoke last. I also counterbalanced the identity of the DDS speaker (experimenter 1 or 2) and the location from which DDS was played (left/right) across trials.

Procedure

The experimental set up and procedure was the same as that used in Chapter 3 with small variations for data collection at Waggy dog crèche. Equipment was set up as described in Chapter 3. Experimenters left the room via door 2. The third experimenter retrieved the dog from its kennel (Redhouse), or from a staff member (Waggy Dog crèche) and entered the experimental room through door 1. The dog was allowed to explore the experimental room for approximately 1 minute, to habituate to the environment with the aim of reducing distraction during the trial. The dog was then guided away from the speakers to wait through door 3 (Redhouse), or around a corner (Waggy dog). Experimenters 1 and 2 entered the room through door 2 and sat in the chairs with the speakers on their laps. The experimenters wore scarves that covered their mouths and did not attempt to move their mouths to simulate speech. They also sat still and did not move apart from to lift and lower their heads according to the eye-gaze condition. They also maintained neutral facial expressions to ensure the dogs did not receive differential facial cues from the experimenters. While the stimulus played, the dog was kept on a short lead to ensure it remained within camera visibility, but the handler did not attempt to redirect the orientation of the dog if it turned or moved within 1m. The handler did not interact with the dog in any way, and looked at the ground throughout. At the end of the stimulus, the handler removed the lead and the dog was allowed to explore freely and approach the experimenters if he/she wished. The dog received no interaction from any experimenters before the end of the trial. At Waggy dog crèche a staff member was present but out of sight of the dog for the stimulus presentation, and did not interact with the dogs if they approached during the trial.

Video Coding

Video recordings of each session were analysed, and during the stimulus presentation, time spent looking towards each speaker was recorded as measured by head direction. During the 1-min off-lead period following the stimulus presentation, time spent in proximity to each speaker was recorded, as measured by the position of the dog's head being within the 1.1 m2 area surrounding each speaker's chair. The period after the dog entered the room, but before the stimulus began was used as a control period (mean duration 5.16 ± 1.89 s). Looking times during this phase were recorded in order to establish whether the dog displayed any preference for one experimenter in particular, or one location (left or right) that may have influenced looking times in the experiment.

Inter Observer Reliability

The primary observer (AB) coded 100% of videos. A second trained observer coded 50% of videos ($N = 26/52$ trials total) and measured looking time at each speaker in each section of the stimulus (control silence, simultaneous 1, DDS only, ADS only, simultaneous 2; $N = 10$ measurements) and time in proximity to each speaker in the minute post-stimulus presentation ($N = 2$ measurements). The primary coder had high agreement with the secondary coder (Spearman's $R > 0.90, p < .001$), indicating the videos had been coded reliably.

A third observer, blind to the hypotheses of the experiment, also coded 23% of the videos ($N = 12/52$ trials total) from clips of the videos where the experimenters were not visible, so the coder could not see the eye-gaze of the speakers, and also with the sound removed, so that they could not tell which type of speech was being played. There was high agreement with the primary coder for looking time ($R = 0.91, p < .001$) and for proximity preference ($R = 0.89, p < .001$).

Statistical Analysis

All data were analysed using IBM SPSS (version 24) with the significance level set at $p < .050$. Attentive and affiliative preference was evaluated using mixed ANOVAs with the fixed within-subject factors speech type (DDS/ADS), and eye-gaze (Direct/Averted). The between-subject factors were DDS identity (experimenter 1/experimenter 2) and DDS location (right/left). After the ANOVA on the total looking time had been completed (Table 4.1), separate ANOVAs were then run for each section of the stimulus (simultaneous; ASD only; DDS only). I would not expect to find a significant main effect of eye-gaze during simultaneous speech, as the experimenters always performed matching eye-gaze during these segments (both direct, or both averted). For this reason eye-gaze has not been included as a factor for simultaneous segments. I applied a more conservative Bonferroni-corrected alpha level to the separate section analyses ($p = .017$) to correct for family-wise error that might have arisen from running multiple tests on the same data set. In these ANOVAs I also included between-subject factors DDS identity (experimenter 1/experimenter 2) and DDS location (right/left). Finally, a single mixed ANOVA was conducted on the proximity to speakers in the minute post-stimulus presentation with speech type as a between-subjects factor. Eye-gaze was not a factor in this analysis as both speakers performed both direct and averted gaze during the trial. All assumptions of these parametric tests were tested and met.

Results

For the following analyses 9 trials were removed; 6 due to equipment failure, 1 due to the dog barking loudly during the trial, making coding impossible, and 2 due to experimenter mistakes during the trial (lifting head at wrong time). This resulted in N=43 for all analyses. During the control silence, there was no significant influence of DDS location or identity, indicating that dogs did not display any a priori preference for either experimenter, or location (Table 4.1).

*Table 4.1. Results from a between-subjects ANOVA for the control period, and a mixed ANOVA (df =1,39) comparing main effects for looking time across all segments of speech. Significant values are presented in bold, * denotes p<.050, *** denotes p<.001.*

	Within-subject Effects F(*p*)			Between-subject effects F(*p*)		
	Speech type	**Eye-gaze**	**Speech type *Eye-Gaze**	**Identity**	**Location**	**Identity *Location**
Control				<0.01 *(.973)*	0.04 *(.846)*	0.04 *(.846)*
Total Looking	**20.21 (<.001)***** η_p^2= 0.34	**5.88 (.020)*** η_p^2= 0.13	0.64 *(.429)*	0.28 *(.598)*	3.54 *(.067)*	0.22 *(.646)*

For total looking time, a mixed ANOVA revealed a significant main effect of speech type (Figure 4.2, Table 4.1) where dogs spent on average longer looking towards a speaker of DDS than one of ADS. The ANOVA also revealed a main effect of eye-gaze (Figure 4.3, Table 4.1), where dogs spent more time looking towards speakers when they had directed eye-gaze than averted eye-gaze. No interaction of speech type and eye-gaze was found and no effects of the between subject factors of DDS speaker identity or location were found (Table 4.1).

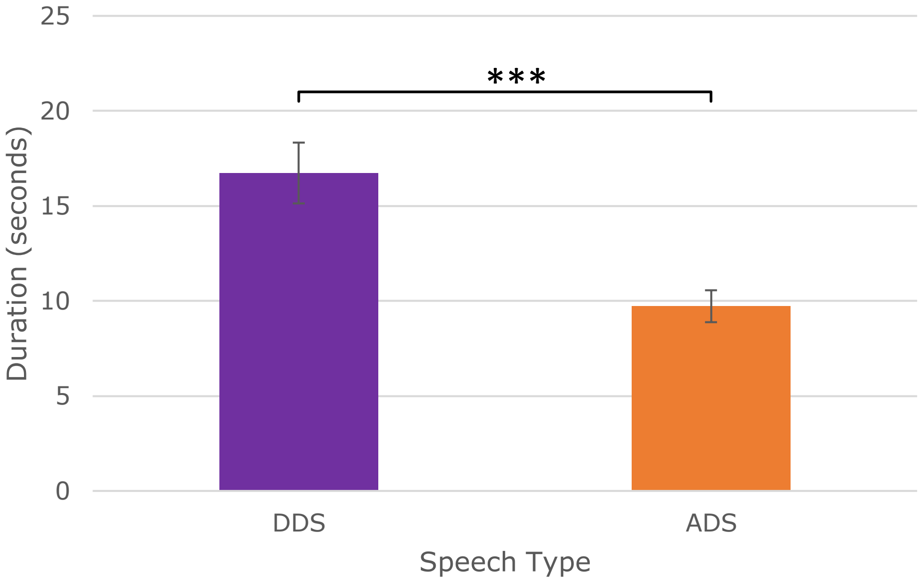


*Figure 4.2 Looking preference towards dog- and adult-directed speech (seconds). Results of a mixed ANOVA revealed a main effect of speech type, where * denotes p<.050, ** denotes p<.010 and *** denotes p<.001. Error bars represent 1 standard error of the mean.*

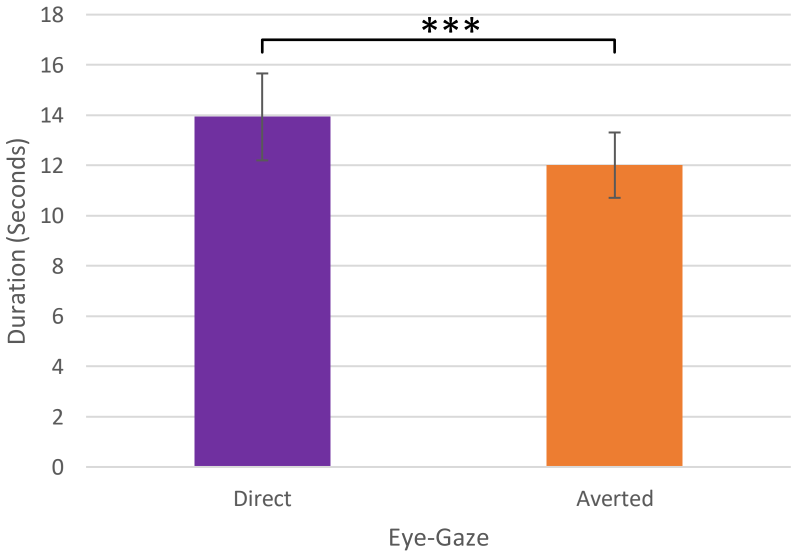


*Figure 4.3 Looking preference towards direct and averted eye-gaze (seconds). Results of a mixed ANOVA revealed a main effect of eye-gaze, where * denotes p<.050, ** denotes p<.010 and *** denotes p<.001. Error bars represent 1 standard error of the mean.*

The three different types of speech segment (simultaneous, DDS only and ADS only) were analysed separately using mixed ANOVAs, with Bonferroni corrected alpha ($p = .017$) to account for multiple comparisons on the same data set. Results of a mixed ANOVA for simultaneous segments revealed a significant main effect of speech type (table 4.2), with dogs preferring to look more towards DDS (mean = 6.00 ± 2.25 seconds) than ADS (mean = 2.80 ±1.32 seconds; table 4.2). For ADS only, a mixed ANOVA revealed no main effect of eye-gaze, and although dogs appeared to look more towards the active speaker (ADS) than DDS (Figure 4.4), this was found to be non-significant (Table 4.2). A mixed ANOVA for DDS only segments revealed a main effect of speech type with dogs preferring to look towards DDS more than ADS (figure 4.4), and a main effect of eye-gaze with dogs preferring to look towards direct eye-gaze than averted. These main effects were qualified by an interaction of speech type and eye-gaze (table 4.2). To explore this interaction effect in the DDS only segment, post-hoc analyses were conducted using paired-samples t-tests with Bonferroni corrected alpha ($p<.004$; figure 4.4). The first t-test revealed that dogs looked significantly more towards DDS with direct eye-gaze than DDS with averted eye-gaze ($t(42) = 4.40$, $p <.001$; Figure 4.4). Secondly, dogs looked significantly more towards DDS with direct eye-gaze than ADS with direct eye-gaze ($t(42) = 6.4$, $p <.001$; Figure 4.4). Dogs also looked significantly more towards DDS with averted gaze than ADS with averted gaze ($t(42) = 3.01$, $p =.004$; Figure 4.4). Finally, there was no significant difference in looking time towards ADS with direct eye-gaze compared to ADS with averted gaze ($t(42) = -0.29$, $p = .771$; Figure 4.4).

For measurement of proximity to the speaker in the minute after the stimulus, a further 3 dogs were removed from analysis, (2 dogs unable to be let off the lead, and 1 due to interruption) resulting in N=40. Time spent in proximity with the speakers was found to violate the assumption of normality (Shapiro-Wilk < .001). A Wilcoxon signed ranks test revealed that there was no difference in the time spent in proximity to either speaker (DDS Median = 0.00 seconds; IQR = 6.75, ADS Median= 4.00, IQR = 7.63; Z=1.33, *p* = .185).

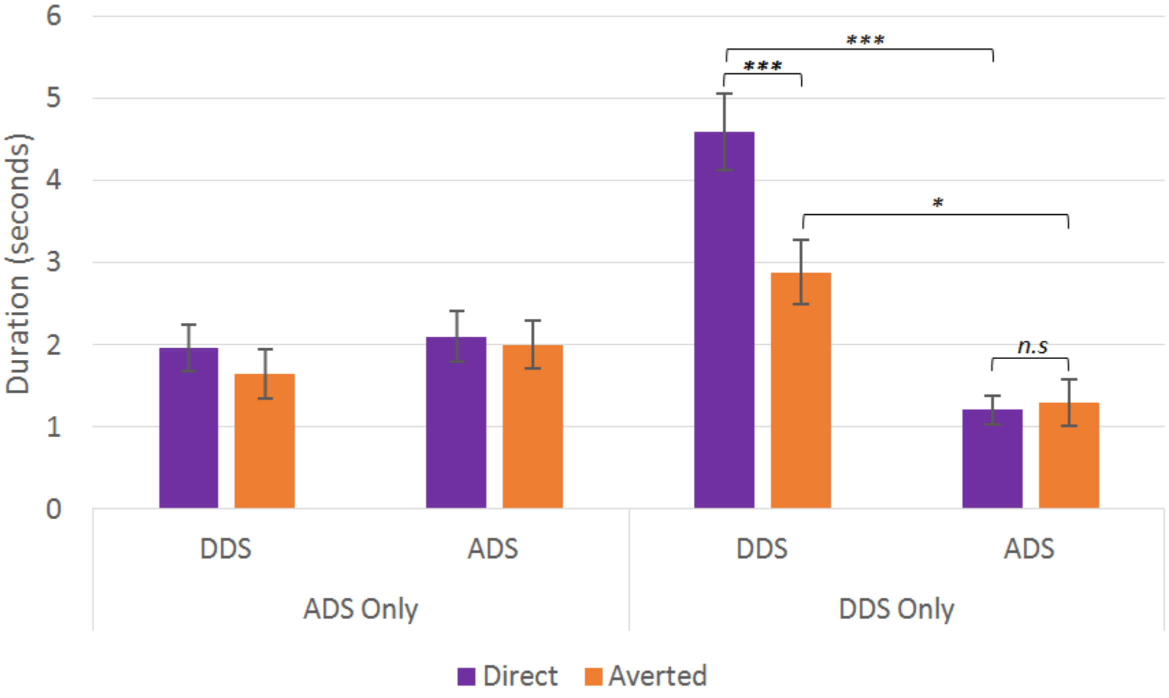


*Figure 4.4. Looking preference towards dog- and adult-directed speech for direct and averted eye-gaze (seconds) in each segment of the stimulus. Results of post-hoc analyses using paired-samples t-tests for the DDS only segment are displayed, where * denotes p<.004 and *** denotes p<.001. Error bars represent 1 standard error of the mean.*

*Table 4.2 Results of mixed ANOVAs (df = 1, 39) for each segment of the trial with Bonferroni corrected alpha (p<.017). Significant effects are marked in bold. ** denotes p<.010, *** denotes p<.001.*

	Within-subject Effects F(*p*)			Between-subject effects F(*p*)		
	Speech type	**Eye-gaze**	**Speech type *Eye-Gaze**	**Identity**	**Location**	**Identity *Location**
Simultaneous	**23.76** **(<.001)***** η_p^2= 0.38			0.04 (*.847*)	0.07 (*.067*)	0.03 (*.866*)
ADS Only	2.46 (*.125*)	0.37 (*.681*)	0.71 (*.404*)	0.37 (*.547*)	3.00 (*.091*)	0.37 (*.547*)
DDS Only	**30.72** **(<.001)***** η_p^2= 0.44	**18.14** **(<.001)***** η_p^2= 0.32	**8.93** **(.005)**** η_p^2= 0.19	0.33 (*.567*)	2.45 (*.127*)	0.19 (*.665*)

Discussion

The results revealed a significant interaction of speech type and eye-gaze during DDS. Post-hoc analysis revealed that dogs preferred to look towards DDS with direct compared to averted eye-gaze. This was not true for ADS. This supports the hypothesis that direct eye-gaze would selectively enhance attention towards DDS, but not ADS, and suggests that eye-gaze provides dogs with an additional cue that the accompanying speech is meant for them. The results also revealed a significant main effect of speech type, with dogs preferring to look towards DDS compared to ADS across the whole stimulus, and for all segments including DDS. This supports the hypothesis and replicates my previous findings discussed in Chapter 3. Further, the results showed that dogs prefer to look towards a human sharing direct eye-gaze with them, compared to one with averted gaze. This supported the hypothesis that eye-gaze would increase attention, and agrees with previous literature that demonstrates that dogs are sensitive to the eye-gaze of humans (Nagasawa et al, 2015, Miklosi et al, 1998, Miklosi et al, 2003). Unlike the preference for DDS over ADS, eye-gaze was not found to significantly influence attention during simultaneous or ADS only speech segments. This could suggest that sound may be the primary cue for dogs when communicating with humans in a naturalistic setting, and that eye-gaze may be less salient compared with speech type for these pets. This is interesting as it suggests that looking at a dog while speaking in DDS enhances the efficacy of the interaction, while the addition of eye-gaze does not influence attention towards ADS.

In terms of proximity to the speakers, this experiment failed to replicate my previous finding from Chapter 3, which demonstrated that dogs preferred to spend time near the speaker of DDS than of ADS. This did not support the hypothesis that dogs would show an affiliative preference for a speaker of DDS. I believe that this result is due to the lack of consistent visual cues indicating affiliation in this study. In my previous experiment, the speakers looked towards the dog at all times and as I have shown, dogs show an attentive preference for human eyes directed towards them, as well as an interaction of eye-gaze with speech type. Previous work has shown the importance of direct eye-gaze for affiliative bonding in particular, for human infants (Kim, Fonagy, Koos, Dorsett & Strathearn, 2014) and in dogs (Nagasawa et al, 2015). It is likely that in my original experiment that consistent directed eye-gaze paired with DDS signalled to the dog that the experimenter was interested in communication and this encouraged later affiliative engagement with the experimenter. In this experiment, the DDS speaker had averted gaze for most of the stimulus presentation period (5/7 speech segments with eye-gaze averted), and this may have created a less compelling signal of interest in affiliation, which in turn led to lower overall motivation in the dogs to interact with the speakers. It is also likely that the alternation of direct and averted gaze was confusing to the dog, and may have led to greater caution in deciding whether or not to approach the speaker. A future experiment where eye-gaze is equally towards and away from the dog, or where one experimenter looks towards while another looks away, would confirm whether affiliative behaviours increase with a higher frequency of shared gaze, as predicted by physiological studies (Nagasawa et al, 2015). These results do, however, suggest that whilst DDS is sufficient to attract the attention of dogs, the combination of directed eye-gaze and DDS may be necessary to affect later affiliation with an unfamiliar individual.

While previous research has shown clearly that dogs are sensitive to eye-gaze cues from humans (Miklosi et al., 1998) this is the first study to find an interaction between eye-gaze and speech type. This suggests that speech type and eye-gaze are two important cues for communicating with dogs, and that eye-gaze appears to selectively enhance the efficacy of DDS. This is consistent with results from studies with infants, such as Haith, Bergman and Moore (1977) who found that while scanning a talking face, infants focus on the eyes more so than when the face is stationary or moving. Although there is some evidence of sensitivity to eye-gaze in other animals (Ferrari, Kohler, Fogassi & Gallese, 2000), to my knowledge this is the only study that assesses the interaction of speech and eye-gaze in non-humans. This study provides evidence against the theory that interpretation of eye-gaze signals as a communicative signal is an adaptation exclusively present in humans (Povinelli & Giambrone, 2001).

From my results it seems that eye-gaze may provide dogs with an additional cue that a human intends to communicate with them. This is reflective of the results found in human infants that suggest eye-gaze is an important ostensive cue that plays a role in natural pedagogy (Kleinke, 1986; Hains & Muir, 1996; Symons, Hains & Muir, 1988; Csibra & Gergely, 2009). Alternatively, it is possible that in unfamiliar speakers, DDS assures the dogs that engaging in direct gaze is safe, whereas with ADS, perhaps engaging in direct gaze is an increased risk, as direct gaze may be interpreted as a threat. It is not possible from these results to conclusively argue in which direction the effect occurs, and aiming to understand a causal link between speech type and eye-gaze could be an interesting topic for future study. In conjunction with previous research, my behavioural results support the idea that eye-gaze is important for dog-human relationships, and dogs' ability to use gaze-cues is likely a fundamental difference between dogs and wolves, who do not engage in gazing behaviours with humans, even following intense socialization from birth (Miklosi et al, 2003). Further, the results of the current study suggest that eye-gaze cues may be particularly important for affiliative bonding with humans, and that the engagement in eye-gaze behaviours may have greater social than attentive benefits, which is likely why I see a decrease in affiliative behaviour in this study, compared with my previous work, where eye-gaze was prolonged.

In conclusion, this study found that direct eye-gaze with humans, selectively enhanced attention towards a speaker of DDS. This interaction of speech type and eye-gaze has not previously been studied, and suggests that engaging in direct gaze with dogs may be beneficial when communicating verbally with them. The lack of affiliative preference for either speaker in this study, in stark comparison to the affiliative preference shown for the DDS in Chapter 3, suggests that the sharing of direct eye-gaze may be particularly important for facilitating social bonds with dogs, in a similar way to that which is observed in human-infant interactions.

Chapter 5: Contextual Barking: A Cross-Species Cue for Social Communication

Abstract

Previous studies have shown that dogs are able to discriminate between barks produced in different contexts and by different individual callers, and humans are able to extract contextual and emotional information from barks produced in various scenarios including play, defence, and solitude. Here I aimed to explore whether dogs *understand* the contextual cues present in barks of conspecifics, by examining whether dogs alter their behaviour appropriately upon hearing barks in play, or in defence. In a within-subject design, 38 dogs heard a short bark bout produced in either a play or defence context, from behind an occluded gateway on their usual walk route. When passing the gate a second time, dogs heard a short bark bout from the other context (play or defence). Attention to the speaker was measured from the coded video recordings during the trial. Dogs did not appear to discriminate between the barks of different contexts, and did not demonstrate any behaviours that suggested they understood the eliciting context of the barks. Using the same stimuli, which varied in acoustic structure with context, humans were able to easily identify the context of the barks. These results suggest that while dogs may be capable of discriminating between contextual barks in some situations, in an ecologically valid paradigm, with short bark bouts, dogs do not alter their behaviour in a way that suggests that they understand the contextual cues. Dogs seem less sensitive than humans to contextual cues in barks, and increased barking in domesticated dogs may have evolved as a means of communication with humans rather than conspecifics.

Introduction

Both domestic dogs, and their close living relatives, wolves, are capable of producing a wide variety of vocalizations. Dogs produce vocalizations frequently, and in a multitude of different contexts (Bradshaw & Nott, 1995; Cohen & Fox, 1976). Wolves on the other hand, while physically able to make a range of sounds similar to that of dogs, vocalize in only very specific contexts (including territorial defence and guarding of food). For wolves, barking is extremely rare (usually only used as a signal for alarm), and howling rather than barking tends to be used to seek attention or social contact from conspecifics (Harrington & Mech, 1979, Klinghammer & Laidlaw, 1979).

It is possible that barking has evolved in domesticated dogs because it facilitates social communication with humans. As highlighted in Chapter 1, modern dogs are highly social and are able to engage in joint attention and cooperation with humans to achieve goal-directed actions (Range & Virani, 2014; Miklosi et al., 2003), and can gain extensive receptive vocabularies using processes similar to those used by humans (Andics et al., 2016). The increase in cross-species sociality in dogs may explain the observed increase in the production of vocalizations. However, it is not clear whether barking has evolved for the purpose of communicating with humans, with conspecifics, or whether it is perhaps a non-functional by-product of domestication (Coppinger & Feinstein, 1991).

The context-specificity of barks may suggest a functional role for barking in social communication. Acoustic properties of barks are observed to be consistent within contexts (Yin, 2002). Yin and McCowan (2004) recorded barks produced in defence, isolation, and play with another dog or human. Following analysis of the acoustic properties of barks in each of these contexts, Yin and McCowan found that defence barks were lower in pitch, longer in duration, and had the most repetitions than the other two types of bark. Isolation and play barks were higher in pitch, and had greater variation in pitch and amplitude. The systematic differences in the acoustic properties of barks from positive and negative contexts have also been mirrored in other dog vocalizations such as growling in play or defence (Farago et al., 2010). The nature of these differences accord with Morton’s (1977) structural-motivational rules, which highlight that across many species of birds and mammals, harsh, low-frequency sounds tend to be used in hostile contexts, and higher-frequency, tonal sounds used when frightened, appeasing, or approaching in an affiliative manner. Morton argues that there is a relationship between the physical structures of sounds and the motivation that elicits their use. These motivational-structure rules may explain the context-specificity observed in dog barks.

Understanding the mechanisms involved in the production of context-specific vocalizations is important, but it is equally important to examine the sensitivity and understanding of listeners to these signals. Human listeners seem to be sensitive to differences in the acoustic structure of human vocalizations and other emotional signals from a very young age. At around 5-7 months of age infants develop an attentional bias towards salient emotional, and social cues, looking more towards fearful visual and auditory stimuli than neutral or happy stimuli (Peltola, Leppanen, Maki & Hietanen, 2009). They are also able discriminate between recordings of their own cries, and cries of another infant (Dondi, Simion & Caltran, 1999). Some people have argued that this is evidence for early empathy in human infants, however while this evidence suggests that they can discriminate between cries of different identities and emotional contexts, it does not indicate that they *understand* the contexts, or the emotion behind the cries. A little later in development, babies are able to *use* emotional cues to inform their behaviour choices. Using a paradigm originally described by Gibson and Walk (1960), infants were placed before a visual cliff apparatus, which is made up of a glass surface above a visible drop, creating the illusion of a cliff or step. In this situation, infants used emotional signalling from their mother to disambiguate the situation (Sorce et al, 1985). Fewer infants attempted to navigate the cliff when the mother showed a fearful or angry facial expression, compared to when she produced a positive expression. When no visual depth was presented, most babies did not look to the mother for a cue, and infants who did look towards the mother hesitated in crossing the cliff in the presence of a fearful face, but crossed regardless. This evidence suggests that emotional signals are useful for regulating infant behaviour, but are most effective in contexts of uncertainty. This evidence is more convincing in suggesting that infants understand emotions and their different contexts, and that they are able to make decisions for their own behaviour, based on communicative signals.

The ability to extract information about the context that elicited certain vocalizations continues to develop into adulthood in humans. Adults can discriminate between different human infant cries including hunger, pain and pleasure (Sagi, 1981), although mothers perform better than non-mothers, and accuracy of discrimination improves after birth, with percentage correct increasing from 48 % correct at 1-day post-partum, to 81% at day 8 (Cismaresco & Montagner, 1990). This suggests that experience of infant vocalizations is beneficial in accurately predicting context. The human ability to extract meaning from vocalizations also extends outside their own species. When recordings of cat solicitation purrs were played to humans, even those with little or no experience with cats judged the purrs recorded when the cats were actively seeking food to be more urgent and unpleasant, compared with purrs recorded in other contexts (McComb, Taylor, Wilson & Charlton, 2009). Acoustic analysis revealed that the solicitation purrs contained a high-frequency voiced component, similar to that of a human infant cry. It was this voiced component that predicted the urgency and pleasantness ratings, and when this component was removed, purrs were considered significantly less urgent. The authors argue that these results suggest companion animals such as cats are able to exploit our inherent sensitivity to vocal signals relating to nurturing offspring.

Cats are not the only domesticated non-human species whose vocalizations can be understood by humans. As discussed in Chapter 1, human participants with varying degrees of dog experience, are able to identify the contexts of barks significantly above chance (Pongracz, Molnar, Miklosi & Csanyi, 2005), and interestingly, when asked to make judgements about the affective state of the dog producing the bark, their evaluations corresponded with the context in which the bark was recorded, and correlated with its acoustic properties. These authors argue that production of contextually specific barks may function as a communicative tool that utilizes the emotional understanding of the listener. Further evidence that the comprehension of contextual barks may be emotionally driven comes from Pongracz et al., (2006) who found that acoustic properties of barks were significantly related to judgements of five emotional states; fear, aggression, despair, playfulness and happiness. Barks lower in pitch with high rates of repetition were categorised as aggressive, and barks higher in pitch, with longer inter-bark intervals were judged to be happy or playful. Together, this evidence suggests that human listeners are sensitive to the acoustic changes of barks produced in different contexts, and goes some way to suggest that barking may provide a communicative signal to humans which conveys information about the environment, or the emotional state of the dog.

While it seems that humans can discriminate between barks produced in multiple contexts, and that they are able to extract meaning from these, it is not clear whether contextually specific barking evolved as a means of communication specifically with humans, or whether dogs also understand and use these barks to communicate with conspecifics. In order to investigate whether dogs can discriminate between barks produced in different contexts, Maros et al., (2008) conducted a series of playback experiments with dogs. Maros et al., used barks produced in defence and isolation while measuring the dogs' heart-rate in a habituation-dishabituation paradigm. Each dog heard three barks of the same context, followed by one bark from a novel context. Barks were from the same individual, and context order was counterbalanced. They found that dogs' heart-rate decreased throughout the first three barks, which indicated habituation. Heart-rate increased in response to the fourth bark which differed in context, indicating that dogs dishabituated in relation to the change in context. This demonstrates that when presented with barks that differ acoustically to reflect a change in context, dogs are able to discriminate between these stimuli on a physiological level. These findings were supported by a behavioural experiment in which looking duration towards the bark was measured (Molnar et al., 2009). They found that repetitions of barks produced in the same context led to decreases in looking duration. Upon hearing a fourth bark from a new context, looking duration increased indicating that the dogs were able to discriminate and that this was reflected in an attentional change. This experiment also found that dogs dishabituated to barks of the same context produced by different individuals. In a control condition, dogs habituated to mechanical sounds, and did not show dishabituation to a subsequent novel mechanical noise. The authors argue that these results suggest that the increase in attentional behaviour for barks is modulated by variations in acoustic properties relating to context, and identity, and not simply stimulus novelty. However, this finding should be interpreted with caution, as lack of dishabituation to the novel mechanical sounds could be due to a lower overall interest in the stimulus category compared to conspecific vocalizations. An alternative explanation of these results is that dogs dishabituated in response to stimulus novelty across contexts and identities, but only do so because the stimulus category is relevant or interesting to them. Nonetheless, this evidence supports the claim that dogs are able to discriminate between acoustically distinct barks, however whether they understand the context associated with the barks is an important question that is not addressed by this research.

While habituation-dishabituation experiments clearly show that dogs can discriminate between barks produced in different contexts (Maros et al, 2008), and by different individual callers (Molnar et al, 2009), in order to determine whether these barks provide meaningful information to listeners about context that elicited them, playback experiments that measure behaviour in an ecologically valid set-up are required. To my knowledge, there is only one previous study that addresses this challenge for barking. Pongracz et al., (2014) recorded barks of dogs left alone in an unfamiliar location (alone barks), and barks produced when an unfamiliar human arrives at the gate to their home (stranger barks). In playback experiments, dogs from multi-dog households heard the alone and stranger barks of unfamiliar dogs, as well as from a familiar resident dog. One minute of recorded barking was played from a speaker at the fenced boundary/gate while the subject dog was alone in his/her garden. They measured duration spent in proximity to the gate and to the house, orientating to the gate and to the house, and frequency of bark production during the playbacks. Their results showed no influence of context on orienting towards the gate or house. In terms of approach behaviour, for unfamiliar but not familiar callers, dogs spent significantly longer near the gate during stranger barks than alone barks. Dogs also barked more frequently in response to stranger barks than alone barks, and this was not influenced by the familiarity of the caller.

Pongracz et al. (2014) argue that these behavioural differences in barking and approach to the gate in response to alone and stranger barks indicate that, particularly for unfamiliar callers, dogs effectively discriminate between barks produced in different contexts, and that these barks influence the behaviour of the listener in an adaptive way. There are, however, other potential explanations for these findings that need exploring. In terms of the difference in bark frequency in the listening dog, the authors argue that this indicates understanding of the higher level of threat in the stranger condition, and may represent recruitment calling. However, use of recruitment calling usually occurs most when a dog is alone (Ortolani et al., 2009), and therefore it would be reasonable to suggest that alone barks are more likely to serve as recruitment barks than stranger barks. The increase in bark frequency in response to stranger barks could simply be due to the greater number of barks in the recordings of stranger barks than alone barks. Without controlling for the number of barks heard, it is difficult to rule out the possibility that increased barking of the listener dogs could be due to mirroring the increased bark frequency in the playback recordings of stranger barks. In terms of the increased duration spent at the gate in response to unfamiliar dog stranger barks, it is unclear why dogs seem to discriminate for the unfamiliar barks, but not for familiar callers. It would be reasonable to argue that a stranger bark produced by a familiar individual should be more salient to the listener (where a collective response to a stranger would be important) compared with that of an unfamiliar caller. This should lead to greater discrimination in familiar, compared with unfamiliar barks; the opposite result to that found by this study. Pongracz et al., argue that their study shows differential behavioural response based on context information, and that this suggests barking is useful in conspecific communication as well as communication with humans. However, this is based on increased bark rate for stranger barks, which could be linked to higher bark frequency in the stimuli, and proximity to the gate in unfamiliar-stranger barks. With no effect of context found for orientation, and no influence on proximity in familiar barks, this conclusion is perhaps too strong. Further research is needed using a range of bark contexts, with ecologically valid set ups, but controlling for these confounding factors, in order to determine whether dogs use conspecific barks to inform their behaviour in a context specific way.

In order to determine whether dogs discriminate between barks in an ecologically valid setting, I conducted playback experiments with dogs on a walk. Barks recorded in play and defence contexts were played from a loudspeaker behind an occluded driveway on the dogs' regular walk route. In the playback experiment, dogs were paused outside the driveway, as one type of bark was played (e.g. play), they then walked on before pausing again outside the drive and hearing the other type of bark (e.g. defence). If dogs understand that the acoustic differences in the barks reflect different situations, or emotional states of the barking dog, I would expect to see a behavioural difference in response to play and defence barks, suggesting that they use these signals to inform their behaviour.

Behavioural and neuroimaging evidence has shown that human adults display enhanced processing of signals that relate to threat, including attentional bias towards fearful faces, compared with neutral faces (Pourtois et al, 2004; Holmes et al, 2005; Viulleumier, 2005). Evidence also suggests that by 7 months old, infants spontaneously look longer towards fearful faces, compared with happy or neutral faces (Nelson & Dolgin, 1985, Kotsoni et al, 2001, Leppanen et al, 2007, Peltola, Leppanen, Maki & Hietanen, 2009). This is thought to be an adaptive response that reflects greater urgency to react to a threat in order to increase chances of survival. Based on this literature, I expected dogs to look longer towards the source of an emotionally salient defence bark than that of a play bark. I also expected more submissive or aggressive behaviours in response to the defence bark, compared with the play bark. If dogs do not display a difference in behavioural response, this may suggest that barking does not provide a meaningful cue to context or emotional state, suggesting perhaps that barking has evolved as a means of communication with humans, rather than with conspecifics. Specifically, I predicted that across dogs, there would be a stronger behavioural response to defence barks than play barks when assessing only the response to the first bark heard. When comparing each dog's response to the play and defence play barks in a within subjects design, I still expected to see a stronger response to defence barks compared to play barks, but it is anticipated that this may be mediated by order effects. Although order effects might be expected, the direction of these effects was difficult to predict: if dogs did not extract meaningful contextual information from the barks, then a stronger response to the first bark, which reveals the presence of an unfamiliar dog might be expected. If however, dogs did understand the eliciting context of conspecific barks, a stronger response to the second bark which indicates a change in context might be expected. Such a response would be in line with Maros et al., (2008; 2009) showing discrimination between bark types based on longer looking and increase in heart rate to the bark from an incongruent context to the preceding barks.

Method

Study site and participants

Dogs were recruited from pet-owners in the village of Newton-Upon Derwent, York, UK, though word of mouth, and a village newsletter (email). All dogs were local to the study site, and regularly walked past the gate-way used in this experiment. 38 dogs took part (17 female, 21 male, Mean Age 5.29 ± 3.42 years).

Stimuli

Dog barks were recorded from 10 dogs (5 male, 5 female) whose owners reported that they produced barks in both play and defence contexts. Recordings were made by experimenters during visits to the dog's home, with the permission of the owner. Barks were recorded by experimenters as the dog owners created play and defence contexts. Play barks were captured from dogs during play with their owner, although the type of play varied between dogs (e.g. chase, tug, fetch). Defence barks were captured during the arrival of someone at the door of the dog's home (door bell, knock on door). Barks were recorded as uncompressed WAV files using a Marantz PMD661 solid-state recorder and Sennheiser microphone. Playback stimuli were created from the recordings of barking bouts in each context. For the playback experiment, a clear segment of the longer recording that contained at least two distinct barks and no background noise was extracted (Mean stimulus duration (start of first bark to end of last bark in sequence); Play = 1.58 ± 0.24 seconds, Defence = 1.58 ± 0.18 seconds). The amplitudes of the barks were modified using Raven Pro 1.4 such that the mean RMS amplitude of each bark was equalized (mean RMS play = 7543 ± 151, defence = 7562± 128).

Acoustic Analysis

Stimuli were analysed using PRAAT (version 6.0.36) with pitch settings 50-1500Hz. For measurements of pitch, voiced calls with a visual pitch line were selected and the mean, minimum and maximum pitch of the segment was extracted using the *Get Pitch* function. The following measurements were taken from the playback stimuli: (i) Number of barks; calculated as the number of discrete bark calls produced in the stimulus; (ii) Bark rate; calculated as the number of barks divided by the duration of barks with higher numbers indicating higher rate (iii) Mean duration of barks; calculated as the mean of all measurements of call duration in the stimulus (iv) Mean inter-bark interval ; calculated as the mean of all measurements of duration between calls in the stimulus (v) Mean pitch; mean pitch measurements from all calls within a stimulus were meaned (vi) pitch modulation; calculated as max pitch-min pitch within each call. Values meaned for all calls within the stimulus. A repeated measures ANOVA revealed a significant influence of bark context on number of barks, bark rate, inter-bark interval, bark duration, and mean pitch (table 5.1).

Table 5.1. A repeated measures ANOVA (df 1,9) revealed a significant influence of bark context on number of barks, bark rate, bark duration, and mean pitch. Significant results are highlighted in bold.

	Mean (SD)		**F**	***p***
	Play	**Defence**		
Number of Barks	2.7 (0.48)	3.7 (0.48)	45.00	**<*.001***
Bark Rate (barks per second)	1.02 (0.20)	1.65 (0.56)	6.99	***.027***
Mean Bark Duration *(seconds)*	0.94 (0.26)	1.22 (0.18)	12.43	***.006***
Mean Inter-Bark Interval *(seconds)*	0.66 (0.22)	0.34 (0.24)	7.41	***.024***
Mean Pitch *(Hz)*	332.40 (49.79)	283.80 (51.99)	7.77	***.021***
Pitch Modulation *(Hz)*	237.30 (133.63)	286.60 (113.15)	1.03	*.336*

This analysis confirms that the acoustic properties of the stimuli used in this experiment agree with those reported in previous literature such as Yin and McCowan (2004), with defence barks having longer duration, lower pitch, and shorter inter-bark intervals.

Design

This experiment used a within-subjects design with each dog hearing barks produced both in play and defence contexts. In order to compare behaviours to hearing a bark of either play or defence, a between-subject comparison of the first bark heard by each dog was also conducted. The barks heard by any one dog were always recorded from the same individual, and dogs heard barks from dogs approximately the same size, or larger than themselves. The order of play and defence barks was counterbalanced, with half of the dogs hearing play first.

Procedure

The experimental set up is displayed in Figure 5.1 below. Wooden boards were used to obscure the view to the driveway. The speaker volume was set in a pre-trial test to a volume at which the barks sounded realistic to a human experimenter in the dog's position in front of the driveway. A sound pressure reading of white noise from the same device was recorded, and this was used to set the volume before each testing session to ensure consistent volume across all trials. A dog was then collected from its house by Experimenter 1 and walked past the boarded gate to familiarize them with the set-up. When Experimenter 1 and the dog reached the start point (1), the experimental trial begun. The experimenter did not interact with the dog during the trial and looked straight ahead when walking, and down at a smartphone when paused at the gate. This was to avoid giving the dog any cue to the bark context during the playback, and a stopwatch on the smartphone was also used to time the playback sections. From the start point (1), Experimenter 1 walked the dog up to the centre of the gate and stopped.

Experimenter 2 played the recorded bark from the iPod when the dog reached the centre of the gate. After 30 seconds Experimenter 1 walked the dog to point 2 and then turned to walk back to the centre of the gate (mean ISI = 74.5 ± 2.55 seconds). Again, when the dog reached the centre of the gate, Experimenter 2 played the second bark. After 30 seconds the dog was walked away from the gate and the experiment ended. The sides of the start point (1) and turn point (2) was counterbalanced such that half of the dogs began the trial on the left, and half the dogs began the trial on the right of the gate.

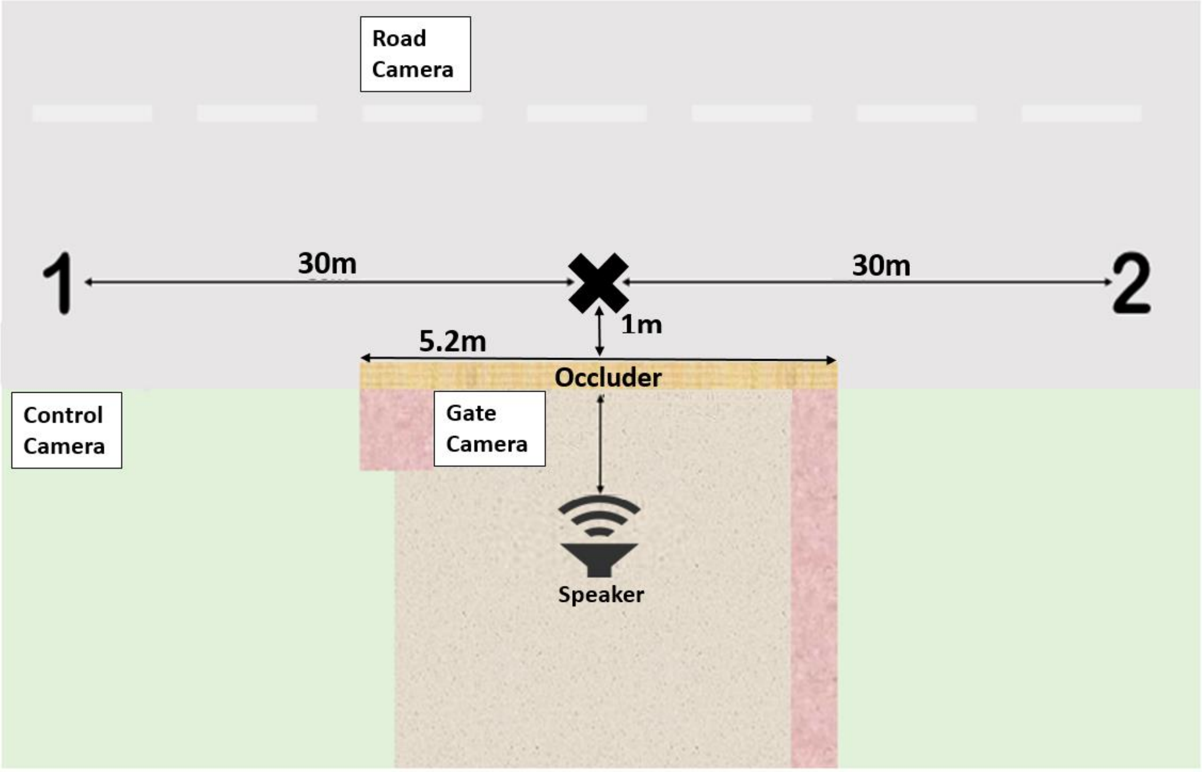


Figure 5.1. Diagram of experimental set up. X represents the position of the dog during the stimulus presentation. Three cameras recorded the dog, one during the control period (1 or 2), and two during the stimulus playback from behind the gate, and from the curb opposite the driveway.

Video Coding

Video recordings of each trial were analysed for time spent looking towards the source of the sound for play and defence barks. The following behaviours were measured from each video: (i) Looking at speaker was measured in terms of the total duration the dog’s head faced the occluded speaker in the period from the start of the stimulus until 30 sec after the start of the stimulus. (ii) Duration of the first look to the speaker; (iii) Latency to look at the speaker was measured as the duration from stimulus onset to the start of the first look to the speaker (iv) Number of looks towards the speaker was measured as the total number of looks towards the speaker (separated by a look away from the speaker of any duration) in the period from the start of the stimulus until 30 sec after the start of the stimulus.

Videos were also coded for behaviours in the control period, and in the 30 second periods following stimulus onset in front of the gateway (Table 5.2).

Table 5.2. Behaviours and their definitions coded in the control period and in the 30 second period during and after stimulus presentation.

	Descriptor
Tail Wag	Movement of the tail left or right at least once
Head Tilt	Rotation of the head left or right on the vertical plane
Lift Ears	Movement of the ears upwards
Attention to Handler	Any attention towards the experimenter holding the lead including looking at, touching, sniffing, or jumping up to the handler
Sniffing	Sniffing the ground
Movement Towards Sound	One or more steps in the direction of the occluded speaker
Movement Away from Sound	One or more steps away from the occluded speaker in any direction
Vocalizations	Any vocalizations including barking, growling, or whining
Scared/Aggressive	Fearful or aggressive behaviours including tucking of the tail, ears down, lip licking, yawning, lunging, hackles up, baring teeth etc.

Inter Observer Reliability

The primary observer (AB) coded 100% of trials (N=31) and measured looking time, latency, number of looks, and duration of first look during each contextual bark (play and defence; N = 8 measurements). A second observer blind to the hypotheses of the experiment also coded 23% of the videos (N= 7/31 trials total) with no sound so that they were not aware which barks were being presented to the dog in the video. There was high agreement with the primary coder for all measures (Spearman's R >.88, $p< .050$), indicating that the videos had been coded reliably.

Statistical Analysis

The behaviours listed in table 5.2 were infrequently observed, particularly when looking for behaviours that were exclusively produced in response to just one of the playback types. I therefore provide descriptive data of the number of dogs who produced each behaviour in the control and post-playback periods. Visualisation of the data revealed that the measures number of looks to the speaker and latency to look to the speaker were distributed over a very limited range (79% of data points were 0-3 looks; latency 85% of data points were 0.5-1 second). Thus there was insufficient variation in these measures for them to be entered into statistical models as continuous variables (even with transforms they would not have been normally distributed) and descriptive statistics only are reported for these measurements.

Assessment of the residuals revealed that the looking time data failed the assumption of normality (Shapiro Wilk >.90, $p <.001$ for all within and between subject measures). Therefore, these data were transformed using a log transform. In order to account for multiple zero values in the dataset, a constant was applied prior to the log transform in order that these data points would be preserved (transformed variable = log10 (raw variable + 1)). Normality of the residuals was reassessed following the transforms, and were found to be normally distributed (Shapiro Wilk test, p>.050 for both measures in both data sets; histograms visually normal). All GLMM analyses were thus run on the transformed looking time data.

Generalized Linear Mixed Models (GLMMs) were conducted separately for the dependent variables total looking time, and duration of first look, with fixed factors *Bark Context, Bark Order* and the interaction between these. All GLMMs were run in IBM SPSS (Version 24) using a normal probability distribution and identity link function. Identity of the stimulus bark was also entered as a random factor to account for repeated sampling of stimulus identity (Waller et al. 2013).

Results

For the following analyses, I removed 3 dogs who did not show any evidence that they had heard the stimulus (e.g. those that did not look towards the source of the sound at all throughout the trial), and 4 dogs whose trial was interrupted by barking from any source other than the stimuli, or pedestrians entering the experimental area. Removing these dogs, resulted in an N=31.

Behavioural Responses

The occurrence of specific behavioural responses (as listed in table 5.2) were evenly distributed over control and post-playback periods, with very few dogs performing any behaviours exclusively in response to one type of playback (see Table 5.3).

Table 5.3. Total number of dogs who perform each behaviour in the control period, and in the 30 seconds following play and defence barks. Number of dogs who perform each behaviour in response to play barks only (not control or defence), and defence barks only (not control or play), are presented in rightmost columns (N=31).

	Control	**Play**	**Defence**	**Play Only**	**Defence Only**
Tail Wag	14	14	13	**0**	**0**
Head Tilt/ Lift Ears	5	9	8	**0**	**0**
Attention to Handler	10	9	10	**2**	**2**
Sniffing	13	17	19	**0**	**1**
Movement Towards Sound	-	13	12	**1**	**1**
Movement Away from Sound	-	14	13	**1**	**0**
Vocalizations	0	1	0	**1**	**0**
Scared/Aggressive	0	0	0	**0**	**0**

Between Subjects: First Bark Only

This analysis compared total looking time, and duration of first look towards only the first bark that each dog heard. Generalized linear mixed models were performed on the transformed looking time data with fixed factor, Bark Context (Play or Defence). Two models were calculated for the target variables total looking time, and duration of first look. These revealed no main effect of bark context for any of the measures (See figure 5.2, table 5.4).

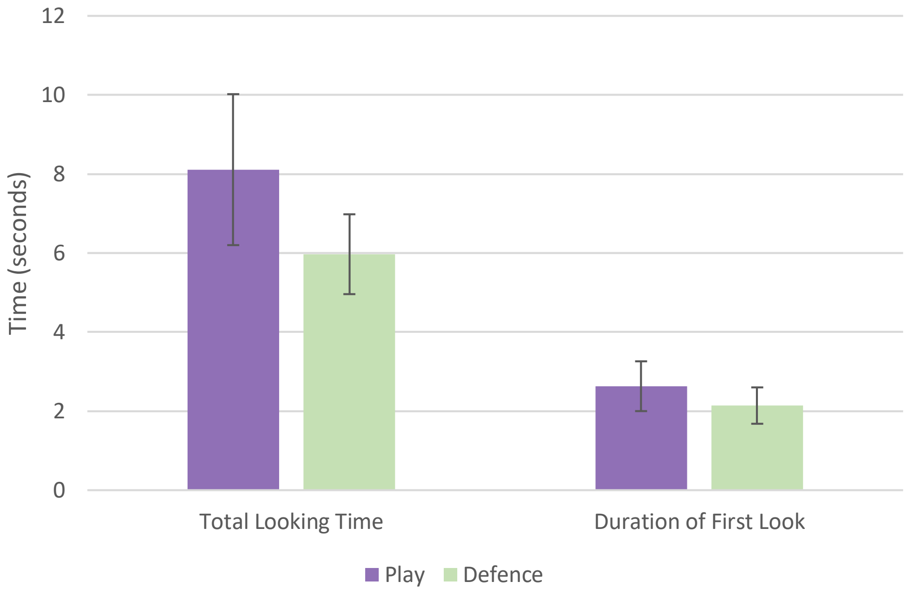


Figure 5.2: Bar graph illustrating (i) the average duration dogs looked to the two types of playback and (ii) the average duration of the first look towards the speaker when played either a play or a defence bark.(first bark only, between subjects) Error bars represent standard error of the mean.

Table 5.4. Results of GLMMs for looking time, and duration of first look. No main effect of bark context was revealed for either measure (N=31).

	F	***Df***	***p***
Looking Time	0.14	*1,29*	*.710*
Duration of first look	0.35	*1,29*	*.558*

The latency to look towards a play bark was 0.46 ± 0.18 seconds, and 0.38 ± 0.15 seconds for a defence bark. The average number of looks towards a play bark was 2.14 ± 0.33, and 2.44 ± 0.35 for a defence bark.

Within Subjects: Two Contextual Barks

The following analysis compared within subject responses to play and defence barks. Generalized linear mixed models revealed no main effects of bark context (Table 5.5; figure 5.3), or bark order (Table 5.5; Figure 5.4), and no interaction between context and order for either looking time measure (Table 5.5). Table 5.6 displays descriptive statistics for the measures Latency to look, and number of looks.

Table 5.5. Results of GLMMs for looking time, and duration of first look. No main effects of bark order or bark context were revealed, and no interaction of order and context was found.

	df	Bark Order F(p)	Bark Context F(*p*)	Order*Context F(*p*)
Looking Time	58	1.44 *(.236)*	0.66 (*.421*)	0.72 (*.401*)
Duration of First Look	58	0.813 (*.371*)	0.40 (*.529*)	1.74 (*.193*)

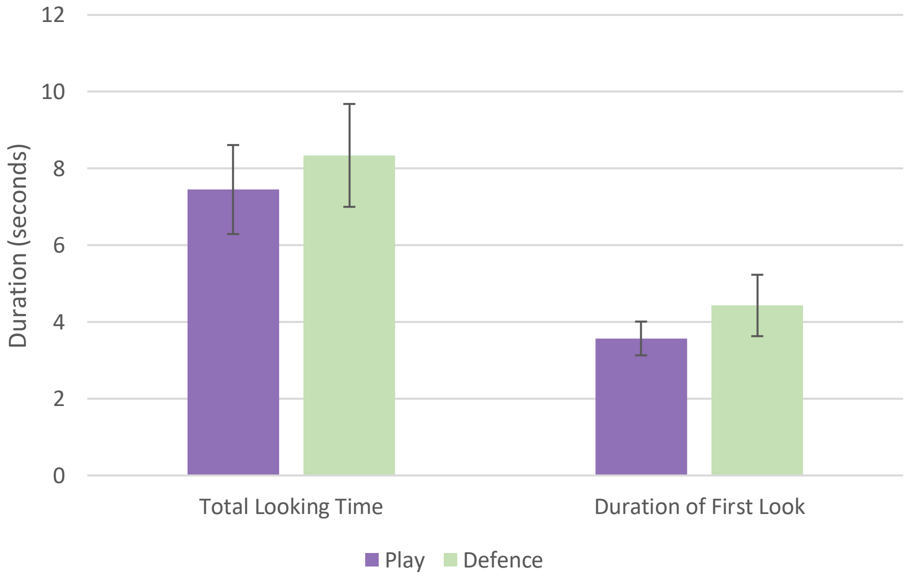


Figure 5.3. Within-subject measures of responses to play and defence barks. Generalized linear mixed models revealed no significant effect of bark context on looking time, or duration of first look. Error bars represent standard error of the mean.

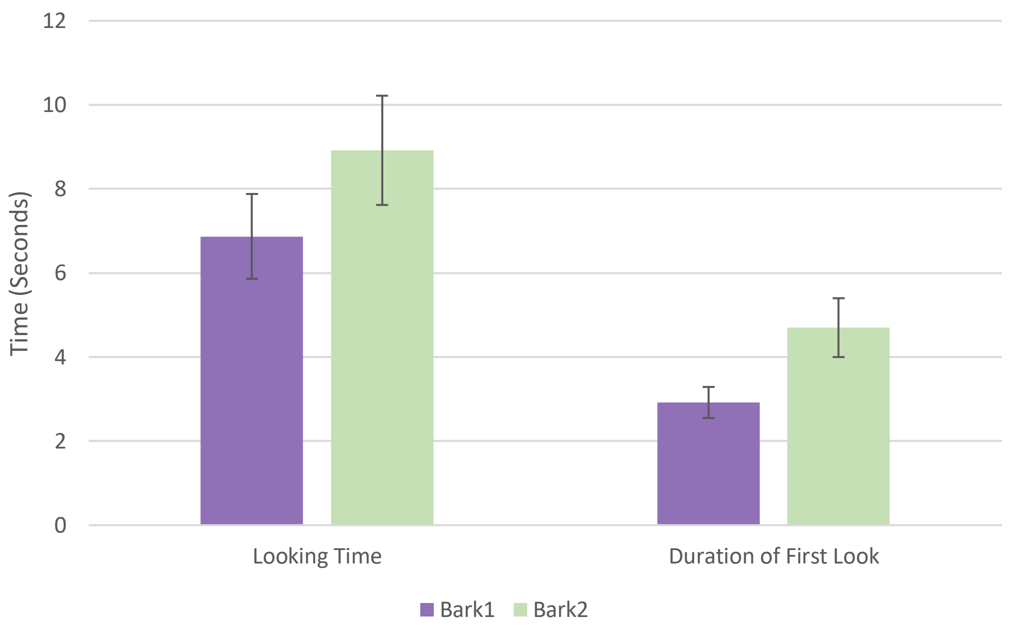


Figure 5.4. Within subject comparison of responses to first and second barks. A generalized linear mixed model revealed no significant effect of bark order on looking time, or duration of first look. There was also no interaction of bark context and bark order. Error bars represent standard error of the mean.

Table 5.6. Means and standard error for measures latency to look (seconds), and number of looks in response to play and defence barks (left) and 1st and 2nd barks (right).

Mean (SE)	**Play**	**Defence**	**1st Bark**	**2nd Bark**
Latency to Look (Seconds)	0.74 (0.29)	0.91 (0.51)	0.43 (0.11)	1.20 (0.54)
Number of Looks	2.26 (0.22)	2.42 (0.25)	2.29 (0.23)	2.38 (0.24)

Dog Playback: Discussion

The results from the above analyses revealed dogs did not behave significantly differently to playbacks of barks produced in different contexts. When comparing between subjects' reactions to the first playback stimulus they heard, there were no differences in looking time or duration of first look for dogs who heard a play bark, compared with those who heard a defence bark. Similarly, when comparing within subjects responses to the first bark, with response to the second bark from a different context, there was no effect of bark context on any response measure. There was also no main effect of bark order, and no interaction between order and context.

One possible reason for this set of null results is that I had an insufficient sample size and thus was underpowered to find effects. In order to evaluate this possibility I conducted a power analysis to assess the number of subjects necessary to detect a medium effect size. I based the effect size in the power analysis on the estimated effect size in the Pongracz experiment, which although used a different methodology is the most relevant previous study. Using G*power (version 3.1.9.2), an a priori power analysis for looking time in my playback experiment indicated that to find a medium effect size based on partial $\eta^2 = 0.2$, with power of 0.95 and an alpha level of 0.05, for the within subject comparison of bark context, 14 subjects would have been required. As I exceeded this with N=31, it is unlikely that the null result in this study is an artefact of insufficient sample size.

These results indicate that dogs do not respond differentially to barks of different context in an ecologically valid setting and showed no signs of discriminating between the two types of barks. This does not support the hypothesis that dogs would show a stronger behavioural response to the emotionally salient defence bark and fails to support the results of Pongracz et al., (2014) and Molnar et al., (2009). Although acoustic analysis showed the play and defence barks used as stimuli to be significantly different on a number of acoustic parameters, there is still a possibility that these differences are hard to perceive and overall these stimuli may have been difficult to discriminate. In order to test this idea, I conducted a follow up experiment in which human participants listened to the same stimuli used in the dog playback experiment, and were tasked with identifying the context of the barks. I predicted that the participants would correctly identify the context of the barks at a level higher than which would be expected by chance. If the humans are unable to perform this task, this would highlight a failure of this particular stimulus set in conveying contextual information to the listener, which could explain the non-significant findings in the dog playback experiments. If, however, the human participants are able to identify the context of these bark stimuli, I can be confident that the null results above are not a reflection of the particular stimuli used. This would support the notion that in an ecologically valid setting, dogs are unable to use contextual information of barks to inform their behaviour.

Experiment 2: Human Playback Experiment

Although the play and defence barks used in this experiment were acoustically different on a number of acoustic parameters, in line with previous literature (e.g. Yin & McCowan, 2004), I wanted to ensure that the lack of significant differences in response to the barks, could not be attributed to ambiguity in this particular stimulus set. I conducted a follow up experiment with human participants in order to determine whether the context of the barks used in this study was detectable.

Methods

Participants

For this experiment, 14 female dog-owners took part (Mean Age = 26.7 ± 9.47).

Stimuli

The audio stimuli used where the same play and defence barks used in the dog playback experiment above, presented in a timed PowerPoint presentation. Participants indicated their responses on a printed answer sheet.

Procedure

Part 1: In order for the humans to complete a task as close to that of the dogs in this study, I first played them one bark by itself and asked them to identify whether the bark was play or defence. They listened to 20 barks in total; play and defence barks from one individual were played consecutively, but the order of play/defence was counterbalanced and randomised between participants, as was the order of individual dog identity.

Part 2: For completeness, I also played participants pairs of barks from the same individual, counterbalanced for order of play and defence, and asked participants to identify which of these two barks they thought was the defence bark (first or second). For each participant the order of the stimuli were randomised in terms of dog identity. The whole experiment took around 10 minutes to complete.

Results

Number of correct responses for Part 1 and Part 2 were counted and percentage accuracy was calculated. When asked whether an individual bark was play or defence, participants responded with 83.57% accuracy. A one-sample t-test (0.50) revealed that this performance was significantly higher than that which would be expected by chance (t (13) = 15.8, *p* <.001). When asked which of two barks was produced in defence, participants were 92.14% accurate. This result was negatively skewed and violated the assumption of normal distribution (Shapiro-Wilk, *p*<.050). Therefore, a one-sample Kolmogorov-Smirnov test was conducted (0.50), and revealed that performance was significantly higher than would be expected by chance (K-S (14) = 0.265, *p* =.009).

Human Playback: Discussion

Human listeners were able to easily identify the context of play and defence barks above the level expected by chance. The results from the human playback experiment clearly show that despite the short length of the barking stimuli selected for this study, human participants are highly capable of understanding the context of play and defence barks. These results support the use of these bark stimuli in the dog playback experiment, and suggest that the lack of significant effects in playbacks with dogs cannot be attributed to a failure of these stimuli to convey contextual information. However, it is important to consider here that while the acoustic properties and the length of the stimuli were sufficient for human listeners to extract contextual information, the human listeners were primed in a way that is not possible during playbacks with animals. The human listeners were told that they would hear a bark, and also asked to identify the context. These instructions alone could have helped the human participants to attend to contextual cues and allowed them to gather the information they required to make their decision in a much shorter time than if they had not been primed. There is evidence that animals are also sensitive to attentional priming, such as improved performance on a visual search task following spatial priming in pigeons (Blough, 1989); importantly, reaction time decreases if a target is primed. Therefore, while human playback results suggest that these stimuli are sufficient in length for discrimination, this does not rule out the possibility that the dogs were not able to attend and react to the contextual cues in such a short time.

Discussion

The results of this study suggest that, when played just one short contextual bark bout, dogs do not show any significant differences in response to barks produced in play or defence. When played a second bark bout from a different context, dogs did not show any change in response that suggested any discrimination of the contextual information. This was true for measures of looking time and duration of first look, as reflected in the results of the GLMMs. There was also no evidence from the behavioural observations that suggested the dogs altered their behaviour in response to the different bark contexts. These results do not support the prediction that dogs would display a stronger response to defence barks. I am confident that result could not be attributed to lack of statistical power or the particular stimuli used in this study as further analysis revealed significant differences in the acoustic properties of the stimuli in line with previous literature, and human participants were able to identify the context of these barks with great accuracy.

On first inspection, these results seem to disagree with previous findings that suggest dogs are able to discriminate between barks produced in different contexts. Molnar et al. (2009) showed clearly that in a dishabituation experiment, dogs show increased attention to a bark of one context following habituation to a bark of a different context. This has also been shown in terms of heart rate increase in response to a contextual change following habituation (Maros et al. 2008). These dishabituation paradigms clearly show that dogs are capable of discriminating between barks produced in different contexts in the same way that they are able to discriminate between identity of the caller (Molnar et al. 2009). Although I did not observe any increase in behavioural response to a second bark of a different context, this is not surprising as the methods were not designed to test discrimination directly and thus were very different to those in previous studies. Firstly, I only played each bark bout once, whereas Molnar et al. (2009) played 3 bark sequences from the same context during habituation before presenting a new context. There was also a larger inter-stimulus-interval in my experiment compared with previous literature. With shorter inter-stimulus-intervals and multiple repetitions, this study may too have observed discrimination of the bark contexts. However, the present study was designed to find out not only whether dogs discriminate between contextual barks in an ecologically valid setting, but also to investigate whether they use information in the barks to influence their behavioural response.

The results of the present study suggest that in a scenario more reflective of everyday life, dogs don't seem able to discriminate between barks produced in different contexts, and perhaps more importantly that they do not appear to *use* this contextual information to inform their behaviour. This highlights that while dogs may be able to discriminate context from acoustic properties in some circumstances, this does not mean that they *understand* context in the same way that humans do. This result is particularly interesting as it may suggest that contextual barking may be uniquely useful for dog-human communication, rather than for communicating contextual cues to conspecifics. It should be noted that the limited range of speaker frequencies may mean that frequencies that would normally be audible to dogs listening to barks could be absent from the stimuli. It is possible that this also reduces the salience of the stimuli to the dogs, but does not influence human perception of the barks as the range of frequencies audible to humans is within the speaker frequency range.

These results also disagree with findings from the previous research of Pongracz et al (2014) who found that, in their own gardens, dogs were able to discriminate between 'stranger' and 'alone barks' of unfamiliar dogs. There are several differences in methodology between these two experiments, which could account for the differences in results. As discussed previously, the stimulus duration of these two experiments is vastly different. With 1 minute of barking stimuli in each context, dogs in Pongracz et al.'s study had far longer to determine the context than the dogs in my study who only heard a few seconds of each bark bout. It is possible in that time, that emotional contagion occurred, with prolonged exposure to the stranger bark, making the listening dogs more likely to react defensively (approaching territory boundaries). The short bark sequences used in my study are more representative of the type of interactions companion dogs have with other dogs in their daily lives, but may not be sufficient for extracting contextual cues. In their evolutionary history, most interactions with conspecifics would have been multimodal in nature with visual signals combined with barking. Whilst its likely dogs are highly adept at responding appropriately to such multimodal cues, these results suggest that the auditory signals in isolation are not sufficient to reliably support an adaptive response. Secondly, the presence of a handler may have mediated any large differences in response for the dogs in my study. When on a lead, most well-socialised dogs are taught that the human will guide their behaviours, and therefore the dog is perhaps less likely to react strongly to stimuli that they may find arousing when alone. In the study by Pongracz et al., dogs were left alone in their garden and were all from multi-dog households, and therefore the absence of their owner, as well as their dog companion, may have left them in a state of higher arousal than the dogs in this study, and therefore the barking stimuli may have simply been more salient to the isolated dogs.

Overall, while previous studies have revealed that dogs are able to discriminate between barks of conspecifics in highly controlled experimental set-ups or in specific circumstances (e.g. responding to unfamiliar but not familiar dogs) the results of this study seem to suggest that when walking with a human, dogs do not respond differentially to barks of another dog produced during play and defence. So, although it seems within the dogs’ capacity to discriminate barks from others based on context, they perhaps don’t do so readily or easily. In contrast, this study and previous literature support the idea that humans are highly adept at discriminating and understanding the context that elicited barks (Pongracz, Molnar, Miklosi & Csanyi, 2005; Molnar, Pongracz, Doka & Miklosi, 2006). Thus, barking may have evolved primarily to communicate with humans, and although dogs may be sensitive to acoustic differences in barks they only seem to use this to inform their behaviour in selective circumstances.

Chapter 6: General Discussion

The results from Chapter 2 supported previous findings that humans produce a special speech type, characterised by elevated pitch and increased pitch modulation when interacting with a dog. The previous studies in this area had mainly focused on DDS produced by female dog owners and had not systematically investigated the acoustic features of DDS produced by males, or non-dog owners. My study highlighted that gender may play a role in the selective increase of pitch in DDS compared to ADS, but suggested that the degree of pitch modulation in DDS as well as the content of DDS does not differ greatly between males and females. I also observed no influence of dog experience on the production of DDS, with both owners and non-owners producing DDS in line with acoustic and content characteristics described in previous literature (e.g. Burnham et al, 1998; 2002). It has been suggested that the elevated pitch in DDS functions to address the emotional needs of the listener, and so I tested whether individuals more able to accurately understand the emotional needs of others would elevate the pitch of their DDS compared to their ADS to a greater degree than individuals less able to understand these needs. However, my study suggested that empathetic ability had no influence on the acoustic properties of DDS, and also did not affect the content of speech. Together, these results go some way in supporting the idea that listener-appropriate speech such as IDS and DDS may be universally produced within western culture, although there are individual differences in how much the pitch is elevated in DDS compared to ADS. However, it is important to note that while this study addressed the gender imbalance in the production literature, the participants were all from similar western cultures and so future cross-cultural studies are required in order to examine whether the acoustic properties of DDS found in this and previous studies are universally produced.

Results from Chapter 3 suggest that naturalistic DDS is important not only for increasing attention to speech but also for increasing approach behaviours in human-dog interactions. In line with infant literature, this finding suggests that DDS functions to increase attention to developmentally relevant information (natural pedagogy) as well as addressing the emotional needs of a listener to improve the affiliative relationship between a human and a dog. This study also highlighted that for adult dogs, both content and prosody of speech are important in facilitating the preference for DDS, and neither prosody nor content alone is sufficient for driving attention or affiliative bonding. Results from Chapter 4 also suggest that direct eye-gaze during speech interactions is important in communication with dogs. It may be that eye-gaze provides a cue to dogs that a communication is meant for them, as is the case with human infants, or alternatively DDS provides assurance that looking towards a human who is engaging in direct eye-gaze is safe, where making eye-contact with an unfamiliar human using ADS presents a higher risk. Together the results from Chapter 3 and Chapter 4 support the claim that dogs are capable of using sophisticated social cues from humans when attending to communicative signals and that these cues play a role in affiliative relationships with humans using similar processes to mother-infant interactions.

Finally, Chapter 5 shows that in an ecologically valid setting, dogs do not respond differentially to play and defence barks produced by conspecifics. While previous studies have shown that dogs are able to discriminate between the acoustic properties of contextual barks (Maros et al, 2008; Molnar et al, 2009), this study suggests that dogs do not use contextual information from barks to inform their behaviour, which suggests perhaps that dogs do not understand the context of the produced bark in the same way as human listeners. It could also be that human listeners are simply more sensitive to the acoustic changes in barks produced in different contexts and under different emotional states, and were, therefore, able to extract contextual information from the short bark bouts that were presented in this experiment, whereas dogs may need longer bark bouts to be able to differentiate them. Human sensitivity to the informational content of dog barks, along with the increased frequency of barking in domestic dogs compared to wolves, may suggest that barking in domestic dogs has evolved as a means of communicating with humans, rather than conspecifics.

The special relationship between dogs and humans provides a unique opportunity to examine the evolution of social communication. Understanding which aspects of communication have evolved through self-domestication, and which capacities have arisen as a result of direct human selection, may be useful for understanding social communication in other animal models. For example, cats have shown some ability to use human communicative signals such as pointing (Miklosi, Pongracz, Topal & Csanyi, 2005), and are able to include infant-like acoustic components in their meows to communicate urgency with humans (McComb, Taylor, Wilson & Charlton, 2009). However, cats are not as skilled as dogs at using pointing behaviours to locate hidden food and do not look towards the faces of humans when presented with an impossible task as readily as dogs (Miklosi et al, 2005). Cats also perform infrequent gaze alternations between a desirable food item and a human, where dogs look frequently between the object and the human face. This gaze alternation is important for maintaining joint attention (Carpenter, Nagell & Tomasello, 1998; Mundy et al, 2007) and is thought to be one of the earliest markers of intentional communication in human infants, emerging at around 9 months. Comparing the cross-species communicative skills of dogs and other species is useful in understanding the origin of communicative traits. Cats also underwent self-domestication for many centuries before human selection, but cats represent an earlier state of domestication than dogs (Bradshaw et al., 1999). Unlike dogs, cats tend to avoid making direct eye-contact with humans and conspecifics, similar to wolves (Miklosi et al, 2005; 2006), and this might highlight that gazing behaviour is the result of a longer period of self-domestication in dogs, compared to other species.

It is interesting to consider the point at which human intervention occurred in the domestication process, and how this may have influenced the subsequent cross-species social skills displayed by these animals. For instance, while the date at which horses began to be domesticated is largely speculative, it is not thought that horses underwent a period of self-domestication before human intervention (Levine, 1999). Horses began to be captured by humans for their meat, and skin, and therefore approaching humans was extremely dangerous for wild horses. Humans then trapped and tamed horses for working, and hunting, eventually selectively breeding horses in captivity. Keeping horses alive in captivity was initially very difficult due to problems feeding un-weaned foals, and an understanding of horse behaviour was required to overcome these difficulties. It is thus suggested that domestication of horses relied on the taming of individual animals with a predisposition to breed in captivity, in stark contrast to the domestication of dogs which was the result of natural selection for tameness and non-aggression. In this sense, domesticated horses are the result of human selection, rather than natural selection and self-domestication, and examining the social communicative abilities of horses and dogs may help to understand which traits could have evolved through human selection, and which traits require self-domestication. For example, horses, like dogs, are able to differentiate between emotional facial expressions (Smith, Proops, Grounds, Wathan & McComb, 2016), but unlike dogs, horses are unable to follow distal momentary pointing cues from humans (Maros, Gacsi & Miklosi, 2008). This may suggest that understanding emotions from facial expressions could be the result of human selection, undergone by both dogs and horses, whereas perhaps the ability to follow subtle social signals, such as distal momentary pointing cues may have evolved through self-domestication in dogs, but not horses.

It is not only comparisons between dogs and wolves, and dogs and other domesticated species that can shed light on the origins of social communication. The huge geographic range of wolves, and later, dogs, has led to a wide variety of wolf-like and dog-like species available for study in various locations all over the world. These populations include free-ranging canids, such as dingoes, which are thought to be genetically similar to early proto-dogs, after self-domestication but before breed diversification by humans (Smith & Litchfield, 2010). Some even describe dingoes as the living 'form' of wild dogs, without the selection pressures imposed by humans on modern dogs. Smith and Litchfield (2009) tested 7 pure Australian dingoes on their ability to follow nine human given cues and found the dingoes to be successful on all nine cues, including the distal momentary pointing cue that dogs, but not wolves, cats or horses, are able to follow. This finding lends its support to the idea that it is the self-domestication process, previous to human intervention, that led to dogs' ability to follow subtle and informative cues from humans, which perhaps explains why wild (wolves) and domesticated (cats/horses) species are unable to utilize these types of cue. The wide variety of comparisons possible to make with domesticated dogs will allow us to build a clearer picture of the origin of social communicative cues in a way that is not possible in other animal models.

The role of the dog is one that has changed drastically over time, and while early dogs, like horses, may have been kept primarily for working, such as hunting, herding livestock, guarding, or pulling loads, in current times all over the world, the majority of dogs are no longer kept for working, but instead are kept most commonly for companionship. This is likely in large part due to the ability of dogs to engage with humans socially, communicatively, and emotionally, in a way that is far superior so any other species. With a growing role in the emotional support of their human guardians (Katz, 2004), and a strong influence on the way the family home is experienced (Power, 2008), dogs in particular are an increasingly popular choice as a family pet. This has created a growing enterprise for training, as many aspects of a dog's natural behaviour, such as barking, digging, and territory marking are no longer compatible with the dog's role in human society. The production and sale of dog training books, CDs, classes, and television programs have multiplied as people endeavour to improve the experience of owning a dog, and also to improve the quality of human-dog relationships (Schebetta, 2009). However, despite the growing need to understand human-dog interaction, there is still relatively little in the scientific literature surrounding human-pet communications, and rather, the majority of dog training programs are based on anecdotal evidence alone. Deeper understanding of dog-human communication will help to educate trainers, owners, and veterinarians about the subtleties of communicating with pets. This will helps us to move towards more scientific, evidence based approaches, with focus on building meaningful relationships using positive reinforcement and motivational approaches, and away from the dominance-based, aversive training techniques that have been governing the dog-training world in previous years. My research goes some way to achieving this goal, particularly the finding that dogs prefer naturalistic DDS, which suggests that DDS should be considered for use by kennel staff at shelters, and vet staff, particularly when interacting with an unfamiliar dog for the first time, where DDS may provide a signal to a frightened dog that these humans are not a threat.

In conclusion, the key results of this research suggest that dogs consider both the prosody and content of DDS in their preference for this special speech type, which is produced by people in Western societies with few individual differences. Dogs are also sensitive to human eye-gaze, and direct eye-gaze selectively enhances this preference for DDS. This highlights that communicating with dogs in an appropriate speech type and using direct eye-contact may enhance the attention of dogs to communicative signals, in a similar way to the use of IDS in communication with infants for natural pedagogy. Finally, dogs do not appear to be as sensitive to contextual information conveyed in the barks of conspecifics, in comparison with human listeners. This is demonstrated by their lack of behavioural response to play and defence barks produced in an ecologically valid setting. This could indicate that the increased prevalence of barking in domesticated dogs could have evolved as a means of communicating with humans, and not conspecifics. Together, these results highlight that dogs show a sophisticated ability to engage in social communication with humans, far superior to that demonstrated by other domesticated and non-domesticated species, and that the study of dogs is invaluable in understanding the evolution of social communicative skills, including the roles of natural, and human selection.

Appendix 1

Examples of DDS and ADS

Female ADS

‘Um I owned a Jack Russel from the age of 11 and I looked after and walked her every day, things like that. And then we got a lab border collie cross when I was about 16, um and we’ve had a lot of cats. We’ve got 2 at the moment Sooty and William’

Female DDS

‘Hello. Hi lily. Hello. Good girl. Good girl. Hello. Good girl. Yeh? Oh aren’t you a lovely dog? oh yes you are lovely. Good girl. What a good girl you are. Yeh? You're very calm. Yes you are. Much more calm than any of my dogs. They'd be all over the place. You're a good girl. Good girl. Do you like that?’

Male ADS

‘I’ve always loved dogs. I’ve been a really strong uh dog person um always wanting one until I was a teenager. Uh I’ve had two dogs which I’ve looked after and done the bulk of the, like, walking and feeding for so I’d say I’m heavily involved. Um I had two rabbits um before as a teenager. Um I can’t remember that much about them and didn’t do that much um looking after them either’

Male DDS

‘Hello. Hi. You're so pretty. Hey. Hi. That little tail. Bless you. You're so pretty. So sweet. My dog would be running riot. Are you cute? *Kiss noise* you wanna come? You look like you wanna play with toys don't you? Bless you. Yeh?’

Appendix 2

Transcripts for DDS and ADS

Parts of scripts underlined represent segments that were removed, and in bold indicate segments of the speech that were repeated in experiment 2 to account for differences in word rate between DDS and ADS

Content-Matched DDS

Experimenter 1: *'Oh you're such a good dog… Yes you are… Come here… Come here… You're such a good dog yes you are… Good dog… You're such a good dog yeah… You're such a good dog.' (10 Seconds).*

'Come on come here… Good dog… Come here good dog… You're such a good dog… yes you are… come here… come here… good dog… yes you are… come on … come here… yeah come on… good dog… good dog.' (15 Seconds).

Experimenter 2: *'Oh you're such a good dog… you gonna come here?... you gonna come? Come on then… come on… let's go out… lets go for a walk… oh you're such a good dog… yes you are.' (10 Seconds).*

'Oh good dog… you gonna come here? Come on then… come on… let's go out for a walk… shall we go for a walk… come on then… come here… lets go for a walk… oh you're such a good dog, yes you are… you're such a good dog.' (15 seconds).

Content-Matched ADS

Experimenter 1: *'So I had a massive meltdown with my lit survey last night… I found out that the topic I was doing didn't have enough literature on… or I couldn't get access to the papers… so I had to email my supervisor' (10 Seconds).*

'I went to the gym yesterday and I joined this new class… but it… we had to do like… 15 minutes of aerobic exercise followed by all these weights and squats and crunches and things… um… and then I made the mistake of going into the pool afterwards and did a couple of lengths but… today I'm really feeling it.' (15 Seconds).

Experimenter 2: *'I went to the cinema last night and saw a really good film, it was really funny I really think you should go and see it… it was a comedy about this … err… this girl who was getting married.' (10 Seconds).*

'I went to the gym yesterday and joined this new fitness class. It was really, really hard. You had to do, like, 15 minutes of aerobic exercise and then sit-ups and weights… ugh, it was just so difficult… and then afterwards I thought I'd cool off in the pool… but I think that was a mistake because I'm really feeling it this morning.' (15 Seconds).

Content-Mismatched DDS

Experimenter 3: *'So (gasp)* I had a massive meltdown with my lit survey last night… (gasp) I couldn't get access to the papers (gasp) and I had to email my supervisor.' (10 Seconds).*

'(Gasp) I went to the gym yesterday and I joined this new class... we had to do… 15 minutes of aerobic exercise… followed by all of these weights (gasp) and squats (gasp) and crunches and things.' (15 Seconds).

Experimenter 4: *'(Gasp) I went to the cinema last night… (gasp) and saw this really good film… It was really funny… and I think you should go and see it (gasp).' (10 Seconds).*

'(Gasp) I went to the gym yesterday and joined this new fitness class. It was really, really hard, you had to do like, (gasp) fifteen minutes of aerobic exercise, and then… I had to cool off in the pool but that was a mistake.' (15 Seconds).

Content-Mismatched ADS

Experimenter 3: *'Oh you're such a good dog… yes you are… come here... come here… you're such a good dog, yes you are… You're such a good dog, yeah… you're such a good dog.* ***Oh you gonna come here? You gonna come? Come on then, come on, let's go out****.' (10 Seconds).*

"Come on, come here… good dog, come here… good dog… you're such a good dog… yes you are… come here… come here… good dog… yes you are… come on, come here… yeah come on… good dog… good dog. ***Oh good dog, you gonna come here? Come on then, come on, let's go for a walk. Shall we go for a walk? Come here, let's go for a walk****." (15 Seconds)*

Experimenter 4: *'Oh good dog, you gonna come here? Come on then, Come on... Let's go for a walk, shall we go for a walk? Come on then, come here... Let's go for a walk, oh you're such a good dog, yes you are, you're such a good dog…* ***come here, good dog****.' (10 Seconds).*

'Oh good dog, you gonna come here? Come on then, come here, let's go for a walk, shall we go for a walk? Come on then, come on, come here… Let's go for a walk, oh you're such a good dog, yes you are… You're such a good dog, ***come here, come here, oh good dog come here you're such a good dog yes you are****.' (15 Seconds).*

References

Anderson, J. R., Montant, M., & Schmitt, D. (1996). Rhesus monkeys fail to use gaze direction as an experimenter-given cue in an object-choice task. *Behavioural Processes*, *37*(1), 47–55. https://doi.org/10.1016/0376-6357(95)00074-7

Anderson, J. R., Sallaberry, P., & Barbier, H. (1995). Use of experimenter-given cues during object-choice tasks by capuchin monkeys. *Animal Behaviour*, *49*(1), 201–208. https://doi.org/10.1016/0003-3472(95)80168-5

Andics, A., Gábor, A., Gácsi, M., Faragó, T., Szabó, D., & Miklósi, Á. (2016). Neural mechanisms for lexical processing in dogs. *Science (New York, N.Y.)*, *353*(6303), 1030–1032. https://doi.org/10.1126/science.aaf3777

Andruski, J. E., Kuhl, P. K., & Hayashi, A. (1999). Point vowels in Japanese mothers' speech to infants and adults. *The Journal of the Acoustical Society of America*, *105*(2), 1095. https://doi.org/10.1121/1.425135

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, *23*(2), 223–229. https://doi.org/10.1016/S0163-6383(01)00037-6

Batliner, A., Biersack, S., & Steidl, S. (2006). The prosody of pet robot directed speech: Evidence from children. In *Proceedings of Speech Prosody* (pp. 1–4).

Belyaev, D. K. (1979). Destabilizing selection as a factor in domestication. *Journal of Heredity*, *70*, 301–308.

Ben-Aderet, T., Gallego-Abenza, M., Reby, D., & Mathevon, N. (2017). Dog-directed speech: why do we use it and do dogs pay attention to it? *Proceedings of the Royal Society of London B: Biological Sciences*, *284*(1846).

Benjamin, A., & Slocombe, K. (2018). 'Who's a good boy?!' Dogs prefer naturalistic dog-directed speech. *Animal Cognition*, *21*(3), 353–364. https://doi.org/10.1007/s10071-018-1172-4

Bentosela, M., Wynne, C. D. L., D'Orazio, M., Elgier, A., & Udell, M. A. R. (2016). Sociability and gazing toward humans in dogs and wolves: Simple behaviors with broad implications. *Journal of the Experimental Analysis of Behavior*, *105*(1), 68–75. https://doi.org/10.1002/jeab.191

Blough, P. M. (1989). Attentional priming and visual search in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *15*(4), 358–365. https://doi.org/10.1037/0097-7403.15.4.358

Boesch, C., Hohmann, G., & Marchant, L. F. (Linda F. (2002). *Behavioural diversity in chimpanzees and bonobos*. Cambridge University Press.

Boitani, L., & Ciucci, P. (1995). Comparative social ecology of feral dogs and wolves. *Ethology Ecology and Evolution*, *7*(1), 49–72. https://doi.org/10.1080/08927014.1995.9522969

Bonanni, R., Cafazzo, S., Valsecchi, P., & Natoli, E. (2010). Effect of affiliative and agonistic relationships on leadership behaviour in free-ranging dogs. *Animal Behaviour*, *79*(5), 981–991. https://doi.org/10.1016/j.anbehav.2010.02.021

Bradshaw, J. W. S., Nott, H. M. . (1995). Social and communication behaviour of companion dogs. In J. Serpell (Ed.), *The domestic dog : its evolution, behaviour, and interactions with people* (p. 268). Cambridge University Press.

Bradshaw, J. W., Horsfield, G. F., Allen, J. A., & Robinson, I. H. (1999). Feral cats: their role in the population dynamics of Felis catus. *Applied Animal Behaviour Science*, *65*(3), 273–283. https://doi.org/10.1016/S0168-1591(99)00086-6

Broesch, T. L., & Bryant, G. A. (2015). Prosody in Infant-Directed Speech Is Similar Across Western and Traditional Cultures. *Journal of Cognition and Development*, *16*(1), 31–43. https://doi.org/10.1080/15248372.2013.833923

Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science (New York, N.Y.)*, *296*(5572), 1435. https://doi.org/10.1126/science.1069587

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*(1), 55–72. https://doi.org/10.1111/j.2044-835X.1991.tb00862.x

Bynon, J. (1968). BERBER NURSERY LANGUAGE. *Transactions of the Philological Society*, *67*(1), 107–161. https://doi.org/10.1111/j.1467-968X.1968.tb01132.x

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4), i–vi, 1-143.

Cismaresco, A., & Montagner, H. (1990). Mothers' discrimination of their neonates' cry in relation to cry acoustics: The first week of life. *Early Child Development and Care*, *65*(1), 3–11. https://doi.org/10.1080/0300443900650102

Cohen, J. A., & Fox, M. W. (1976). Vocalizations in wild canids and possible effects of domestication. *Behavioural Processes*, *1*(1), 77–92. https://doi.org/10.1016/0376-6357(76)90008-5

Cooper, R. P., & Aslin, R. N. (1990). Preference for Infant-directed Speech in the First Month after Birth. *Child Development*, *61*(5), 1584–1595. https://doi.org/10.1111/j.1467-8624.1990.tb02885.x

Coppinger, R., & Coppinger, L. (2001). *Dogs : a startling new understanding of canine origin, behavior, and evolution*. Scribner.

Coppinger, R., & Feinstein, M. (1991). `Hark! Hark! The dogs do bark...' and bark and bark. *Smithsonian*, *21*(10), 119.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*(1), 28–30. https://doi.org/10.1038/nn1611

Dickstein, S., Thompson, R. A., Estes, D., Malkin, C., & Lamb, M. E. (1984). Social referencing and the security of attachment. *Infant Behavior and Development*, *7*(4), 507–516. https://doi.org/10.1016/S0163-6383(84)80009-0

Dondi, M., Simion, F., & Caltran, G. (1999). Can newborns discriminate between their own cry and the cry of another newborn infant? *Developmental Psychology*, *35*(2), 418–426. https://doi.org/10.1037/0012-1649.35.2.418

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cognition*, *6*(5), 509–540. https://doi.org/10.1080/135062899394920

Evans, C. S., Evans, L., & Marler, P. (1993). On the meaning of alarm calls: functional reference in an avian vocal system. *Animal Behaviour*, *46*(1), 23–38. https://doi.org/10.1006/ANBE.1993.1158

Faragó, T., Pongrácz, P., Range, F., Virányi, Z., & Miklósi, Á. (2010). 'The bone is mine': affective and referential aspects of dog growls. *Animal Behaviour*, *79*(4), 917–925. https://doi.org/10.1016/J.ANBEHAV.2010.01.005

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–9605. https://doi.org/10.1073/pnas.152159999

Ferguson, C. A. (1964). Baby Talk in Six Languages. *American Anthropologist*, *66*(6_PART2), 103–114. https://doi.org/10.1525/aa.1964.66.suppl_3.02a00060

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*, 477–501.

Ferrari, P. F., Kohler, E., Fogassi, L., & Gallese, V. (2000). The ability to follow eye gaze and its emergence during development in macaque monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(25), 13997–14002. https://doi.org/10.1073/pnas.250241197

Fisher, C., & Tokura, H. (1996). Acoustic Cues to Grammatical Structure in Infant-Directed Speech: Cross-Linguistic Evidence. *Child Development*, *67*(6), 3192–3218. https://doi.org/10.1111/j.1467-8624.1996.tb01909.x

Fox, M. W. (1971). *Behaviour of Wolves, Dogs and Related Canids by Michael W. Fox (1984-06-03): Amazon.co.uk: Michael W. Fox: Books*. New York, NY: Harper & Row.

Furuichi, T. (1997). Agonistic Interactions and Matrifocal Dominance Rank of Wild Bonobos (Pan paniscus) at Wamba. *International Journal of Primatology*, *18*(6), 855–875. https://doi.org/10.1023/A:1026327627943

Furuichi, T. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthropology: Issues, News, and Reviews*, *20*(4), 131–142. https://doi.org/10.1002/evan.20308

Furuichi, T., Connor, R., & Hashimoto, C. (2014). Non-conceptive conceptive Sexual Interactions in Monkeys, Apes, and Dolphins (pp. 385–408). Springer, Tokyo. https://doi.org/10.1007/978-4-431-54523-1_20

Furuichi, T., Idani, G., Ihobe, H., Kuroda, S., Kitamura, K., Mori, A., … Kano, T. (1998). Population Dynamics of Wild Bonobos (Pan paniscus) at Wamba. *International Journal of Primatology*, *19*(6), 1029–1043. https://doi.org/10.1023/A:1020326304074

Gibson, E. J., & Walk, R. D. (1960). The "Visual Cliff", *202*(4), 64–71. https://doi.org/10.2307/24940447

Goldberg, M. A., Hahn, P. F., Saini, S., Cohen, M. S., Reimer, P., Brady, T. J., & Mueller, P. R. (1993). Value of T1 and T2 relaxation times from echoplanar MR imaging in the characterization of focal hepatic lesions. *AJR. American Journal of Roentgenology*, *160*(5), 1011–1017. https://doi.org/10.2214/ajr.160.5.8470568

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual Following and Pattern Discrimination of Face-like Stimuli by Newborn Infants. *Pediatrics*, *56*(4).

Hains, S. M. J., & Muir, D. W. (1996). Infant Sensitivity to Adult Eye Direction. *Child Development*, *67*(5), 1940–1951. https://doi.org/10.1111/j.1467-8624.1996.tb01836.x

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science (New York, N.Y.)*, *198*(4319), 853–855.

Hare, B., Brown, M., Williamson, C., & Tomasello, M. (2002). The Domestication of Social Cognition in Dogs. *Science*, *298*(5598), 1634–1636. https://doi.org/10.1126/science.1072702

Hare, B., & Tomasello, M. (2005). Human-like social skills in dogs? *Trends in Cognitive Sciences*, *9*(9), 439–444. https://doi.org/10.1016/j.tics.2005.07.003

Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573–585. https://doi.org/10.1016/j.anbehav.2011.12.007

Harrington, F. H., & Mech, L. D. (1979). Wolf Howling And Its Role In Territory Maintenance. *Behaviour*, *68*(3–4), 207–249. https://doi.org/10.1163/156853979X00322

Haxby, J., Hoffman, E., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Hemmer, H. (1990). *Domestication : the decline of environmental appreciation*. Cambridge University Press.

Henning, A., Striano, T., & Lieven, E. V. M. (2005). Maternal speech to infants at 1 and 3 months of age. *Infant Behavior and Development*, *28*(4), 519–536. https://doi.org/10.1016/J.INFBEH.2005.06.001

Hirsh-Pasek, K., & Treiman, R. (1982). Doggerel: Motherese in a new context. *Journal of Child Language*, *9*(1), 229–237.

Holmes, A., Winston, J. S., & Eimer, M. (2005). The role of spatial frequency information for ERP components sensitive to faces and emotional facial expression. *Cognitive Brain Research*, *25*(2), 508–520. https://doi.org/10.1016/J.COGBRAINRES.2005.08.003

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6

Joslin, P. W. B. (1967). Movements and Home Sites of Timber Wolves in Alǵonquin Park. *American Zoologist*, *7*(2), 279–288. https://doi.org/10.1093/icb/7.2.279

Kaplan, P. S., Danko, C. M., Kalinka, C. J., & Cejka, A. M. (2012). A developmental decline in the learning-promoting effects of infant-directed speech for infants of mothers with chronically elevated symptoms of depression. *Infant Behavior & Development*, *35*(3), 369–379. https://doi.org/10.1016/j.infbeh.2012.02.009

Kaplan, P. S., Goldstein, M. H., Huckeby, E. R., Owren, M. J., & Cooper, R. P. (1995). Dishabituation of visual attention by infant- versus adult-directed speech: Effects of frequency modulation and spectral composition. *Infant Behavior and Development*, *18*(2), 209–223. https://doi.org/10.1016/0163-6383(95)90050-0

Kaplan, P. S., Jung, P. C., Ryther, J. S., Zarlengo-Strouse, P., Zar-lengo-Strouse, P., Bradley, J., … Kaplan, S. (1996). Infant-Directed Versus Adult-Directed Speech as Signals for Faces We thank Cynthia Angel, Laura Dolph, and Debra Palmer for coding videotapes. *Cooper & Aslin*, *32*(5), 880–891.

Katz, J. (2004). *The new work of dogs: Tending to life, love, and family*. Villard: New York.

Kaye, K. (1980). Why we don't talk 'baby talk' to babies. *Journal of Child Language*, *7*(03), 489–507. https://doi.org/10.1017/S0305000900002804

Kim, S., Fonagy, P., Koos, O., Dorsett, K., & Strathearn, L. (2014). Maternal oxytocin response predicts mother-to-infant gaze. *Brain Research*, *1580*, 133–142. https://doi.org/10.1016/J.BRAINRES.2013.10.050

Kitamura, C., & Burnham, D. (2003). Pitch and Communicative Intent in Mother's Speech: Adjustments for Age and Sex in the First Year. *Infancy*, *4*(1), 85–110. https://doi.org/10.1207/S15327078IN0401_5

Kleinke, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, *100*(1), 78–100. https://doi.org/10.1037/0033-2909.100.1.78

Klinghammer, E., & Laidlaw, L. (1979). Analysis of 23 months of daily howl records in a captive grey wolf pack (Canis lupus). In E. Klinghammer (Ed.), *The Behavior and Ecology of Wolves* (pp. 153–181). New York: Garland STPM Press.

Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical Perception of Facial Expressions by 7-Month-Old Infants. *Perception*, *30*(9), 1115–1125. https://doi.org/10.1068/p3155

Kruska, D. C. T. (2005). On the evolutionary significance of encephalization in some eutherian mammals: effects of adaptive radiation, domestication, and feralization. *Brain, Behavior and Evolution*, *65*(2), 73–108. https://doi.org/10.1159/000082979

Kubinyi, E., Viranyi, Z., & Miklosi, A. (2007). Comparative Social Cognition: From wolf and dog to humans. *Comparative Cognition & Behavior Reviews*, *2*, 26–46. https://doi.org/10.3819/ccbr.2008.20002

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V, Ryskin, V. L., … Lacerda, F. (1997). Cross-language analysis of phonetic units in language. *Science*, *227*, 684–686.

Kukekova, A. V., Nelson, J., Kuchtey, R. W., Lowe, J. K., Johnson, J. L., Ostrander, E. A., … Acland, G. M. (2006). Linkage Mapping of Canine Rod Cone Dysplasia Type 2 (*rcd2*) to *CFA7* , the Canine Orthologue of Human 1q32. *Investigative Opthalmology & Visual Science*, *47*(3), 1210. https://doi.org/10.1167/iovs.05-0861

Kukekova, A. V., Trut, L. N., Chase, K., Kharlamova, A. V., Johnson, J. L., Temnykh, S. V., … Lark, K. G. (2011). Mapping Loci for Fox Domestication: Deconstruction/Reconstruction of a Behavioral Phenotype. *Behavior Genetics*, *41*(4), 593–606. https://doi.org/10.1007/s10519-010-9418-1

Künzl, C., & Sachser, N. (1999). The Behavioral Endocrinology of Domestication: A Comparison between the Domestic Guinea Pig (Cavia apereaf.porcellus) and Its Wild Ancestor, the Cavy (Cavia aperea). *Hormones and Behavior*, *35*(1), 28–37. https://doi.org/10.1006/HBEH.1998.1493

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP Study of Emotional Face Processing in the Adult and Infant Brain. *Child Development*, *78*(1), 232–245. https://doi.org/10.1111/j.1467-8624.2007.00994.x

Levin, H., & Hunter, W. a. (1982). Children's Use of a Social Speech Register: Age and Sex Differences. *Journal of Language and Social Psychology*, *1*(1), 63–72. https://doi.org/10.1177/0261927X8200100105

Levine, M. A. (1999). Botai and the origins of horse domestication. *Journal of Anthropological Archaeology*, *18*(1), 29–78. https://doi.org/10.1006/jaar.1998.0332

Lieberman, D. E., Carlo, J., Ponce de León, M., & Zollikofer, C. P. E. (2007). A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. *Journal of Human Evolution*, *52*(6), 647–662. https://doi.org/10.1016/J.JHEVOL.2006.12.005

MacDonald, D. W., & Carr, G. M. (2016). Variation in dog society: Between resource dispersion and social fl ux. In J. Serpell (Ed.), *The Domestic Dog* (pp. 319–341). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781139161800.016

Maros, K., Gácsi, M., & Miklósi, Á. (2008). Comprehension of human pointing gestures in horses (Equus caballus). *Animal Cognition*, *11*(3), 457–466. https://doi.org/10.1007/s10071-008-0136-5

Martin, A., Schatz, T., Versteegh, M., Miyazawa, K., Mazuka, R., Dupoux, E., & Cristia, A. (2015). Mothers Speak Less Clearly to Infants Than to Adults: A Comprehensive Test of the Hyperarticulation Hypothesis. *Psychological Science*, *26*(3), 341–347. https://doi.org/10.1177/0956797614562453

McComb, K., Taylor, A. M., Wilson, C., & Charlton, B. D. (2009). The cry embedded within the purr. *Current Biology*, *19*(13), R507–R508. https://doi.org/10.1016/J.CUB.2009.05.033

Mcleod, P., Gadbois, S., & Fentress, J. C. (1996). The relation between urinary cortisol levels and social behavior in captive timber wolves Mummichogs View project Wild canids and canines in general View project. *Article in Canadian Journal of Zoology*. https://doi.org/10.1139/z96-026

McLeod, P. J. (1990). Infanticide by female wolves. *Canadian Journal of Zoology*, *68*(2), 402–404. https://doi.org/10.1139/z90-058

Mech, D. L., Adams, L. G., Meier, T. J., Burch, J. W., & Dale, B. W. (1998). *The Wolves of Denali*. University of Minnesota Press. https://doi.org/papers://A270C103-A120-4E61-B0FE-19A2B90778C5/Paper/p3074

Mech, L. D. (1994). Buffer Zones of Territories of Gray Wolves as Regions of Intraspecific Strife. *Journal of Mammalogy*, *75*(1), 199–202. https://doi.org/10.2307/1382251

Miklósi, Á., Kubinyi, E., Topál, J., Gácsi, M., Virányi, Z., & Csányi, V. (2003). A Simple Reason for a Big Difference. *Current Biology*, *13*(9), 763–766. https://doi.org/10.1016/S0960-9822(03)00263-X

Miklösi, Á., Polgárdi, R., Topál, J., & Csányi, V. (1998). Use of experimenter-given cues in dogs. *Animal Cognition*, *1*(2), 113–121. https://doi.org/10.1007/s100710050016

Miklosi, A., Pongracz, P., Lakatos, G., Topal, J., & Vilmos, C. (2005). A comparative study of the use of visual communicative signals in interactions between dogs (canis familiaris) and humans and cats (felis catus) and humans. *Journal of Comparative Psychology*, *119*(2), 179–186.

Mitchell, R. W. (2001). Americans' Talk to Dogs: Similarities and Differences With Talk to Infants. *Research on Language & Social Interaction*, *34*(2), 183–210. https://doi.org/10.1207/S15327973RLSI34-2_2

Mitchell, R. W., & Edmonson, E. (1999). Functions of Repetitive Talk to Dogs during Play: Control, Conversation, or Planning? *Society & Animals*, *7*(1), 55–81. https://doi.org/10.1163/156853099X00167

Molnár, C., Pongrácz, P., Dóka, A., & Miklósi, Á. (2006). Can humans discriminate between dogs on the base of the acoustic parameters of barks? *Behavioural Processes*, *73*(1), 76–83. https://doi.org/10.1016/J.BEPROC.2006.03.014

Molnár, C., Pongrácz, P., Faragó, T., Dóka, A., & Miklósi, Á. (2009). Dogs discriminate between barks: The effect of context and identity of the caller. *Behavioural Processes*, *82*(2), 198–201. https://doi.org/10.1016/j.beproc.2009.06.011

Morey, D. F. (2006). Burying key evidence: the social bond between dogs and people. *Journal of Archaeological Science*, *33*(2), 158–175. https://doi.org/10.1016/J.JAS.2005.07.009

Morton, E. S. (1977). On the Occurrence and Significance of Motivation-Structural Rules in Some Bird and Mammal Sounds. *The American Naturalist*, *111*(981), 855–869. https://doi.org/10.1086/283219

Muller, M. N., Kahlenberg, S. M., Emery Thompson, M., & Wrangham, R. W. (2007). Male coercion and the costs of promiscuous mating for female chimpanzees. *Proceedings. Biological Sciences*, *274*(1612), 1009–1014. https://doi.org/10.1098/rspb.2006.0206

Mundy, P., Block, J., Delgado, C., Pomares, Y., Van Hecke, A. V., & Parlade, M. V. (2007). Individual differences and the development of joint attention in infancy. *Child Development*, *78*(3), 938–954. https://doi.org/10.1111/j.1467-8624.2007.01042.x

Nagasawa, M., Mitsui, S., En, S., Ohtani, N., Ohta, M., Sakuma, Y., … Kikusui, T. (2015). Social evolution. Oxytocin-gaze positive loop and the coevolution of human-dog bonds. *Science (New York, N.Y.)*, *348*(6232), 333–336. https://doi.org/10.1126/science.1261022

Nagasawa, M., Okabe, S., Mogi, K., & Kikusui, T. (2012). Oxytocin and mutual communication in mother-infant bonding. *Frontiers in Human Neuroscience*, *6*, 31. https://doi.org/10.3389/fnhum.2012.00031

Nelson, C. A., & Dolgin, K. G. (1985). The Generalized Discrimination of Facial Expressions by Seven-Month-Old Infants. *Child Development*, *56*(1), 58. https://doi.org/10.2307/1130173

Newman, R. S., & Hussain, I. (2006). Changes in Preference for Infant-Directed Speech in Low and Moderate Noise by 4.5- to 13-Month-Olds. *Infancy*, *10*(1), 61–76. https://doi.org/10.1207/s15327078in1001_4

Ortolani, A., Vernooij, H., & Coppinger, R. (2009). Ethiopian village dogs: Behavioural responses to a stranger's approach. *Applied Animal Behaviour Science*, *119*(3–4), 210–218. https://doi.org/10.1016/j.applanim.2009.03.011

Pal, S. ., Ghosh, B., & Roy, S. (1999). Inter- and intra-sexual behaviour of free-ranging dogs (Canis familiaris). *Applied Animal Behaviour Science*, *62*(2–3), 267–278. https://doi.org/10.1016/S0168-1591(98)00220-2

Peltola, M. J., Leppänen, J. M., Mäki, S., & Hietanen, J. K. (2009). Emergence of enhanced attention to fearful faces between 5 and 7 months of age. *Social Cognitive and Affective Neuroscience*, *4*(2), 134–142. https://doi.org/10.1093/scan/nsn046

Peter D . Eimas , Einar R . Siqueland, P. J. and J. V. (2016). Speech Perception in Infants. *American Association for the Advancement of Science*, *171*(3968), 303–306.

Phillips, J. R. (1973). Syntax and Vocabulary of Mothers' Speech to Young Children: Age and Sex Comparisons. *Child Development*, *44*(1), 182. https://doi.org/10.2307/1127699

Pongrácz, P., Molnár, C., Miklósi, Á., & Csányi, V. (2005). Human Listeners Are Able to Classify Dog (Canis familiaris) Barks Recorded in Different Situations. *Journal of Comparative Psychology*, *119*(2), 136–144. https://doi.org/10.1037/0735-7036.119.2.136

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological Correlates of Rapid Spatial Orienting Towards Fearful Faces. *Cerebral Cortex*, *14*(6), 619–633. https://doi.org/10.1093/cercor/bhh023

Povinelli, D. J., & Giambrone, S. (2001). Reasoning about Beliefs: A Human Specialization? *Child Development*, *72*(3), 691–695. https://doi.org/10.1111/1467-8624.00307

Power, E. (2008). Furry Families: Making a Human-Dog Family through Home. *Social and Cultural Geography*, *9*(5), 535–555.

Prato-Previde, E., Fallani, G., & Valsecchi, P. (2006). Gender Differences in Owners Interacting with Pet Dogs: An Observational Study. *Ethology*, *112*(1), 64–73. https://doi.org/10.1111/j.1439-0310.2006.01123.x

Pusey, A., Murray, C., Wallauer, W., Wilson, M., Wroblewski, E., & Goodall, J. (2008). Severe Aggression Among Female Pan troglodytes schweinfurthii at Gombe National Park, Tanzania. *International Journal of Primatology*, *29*(4), 949–973. https://doi.org/10.1007/s10764-008-9281-6

Range, F., & Virányi, Z. (2014). Wolves Are Better Imitators of Conspecifics than Dogs. *PLoS ONE*, *9*(1), e86559. https://doi.org/10.1371/journal.pone.0086559

Remick, H., Raffler-Engel, W. von, & Lebrun, Y. (1976). Baby talk and infant speech.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science Research Library Pg*, *274*(5294).

Sagi, A. (1981). Mothers' and non-mothers' identification of infant cries. *Infant Behavior and Development*, *4*, 37–40. https://doi.org/10.1016/S0163-6383(81)80005-7

Schachner, A., & Hannon, E. E. (2011). Infant-directed speech drives social preferences in 5-month-old infants. *Developmental Psychology*, *47*(1), 19–25. https://doi.org/10.1037/a0020740

Schebetta, L. J. (2009). Mythologies and Commodifications of Dominion in The Dog Whisperer with Cesar Millan. *Journal for Critical Animal Studies*, *7*(1), 107–131.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*(4), 1070–1094. https://doi.org/10.1016/S0003-3472(80)80097-2

Smith, A. V., Proops, L., Grounds, K., Wathan, J., & McComb, K. (2016). Functionally relevant responses to human facial expressions of emotion in the domestic horse (Equus caballus). *Biology Letters*, *12*(2), 20150907. https://doi.org/10.1098/rsbl.2015.0907

Smith, B. P., & Litchfield, C. A. (2010). Dingoes (Canis dingo) can use human social cues to locate hidden food. *Animal Cognition*, *13*(2), 367–376. https://doi.org/10.1007/s10071-009-0287-z

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of Child Language*, *4*(01), 1–22. https://doi.org/10.1017/S0305000900000453

Sorce, J. F., Emde, R. N., Campos, J., & Klinnert, M. D. (1985). Maternal Emotional Signaling: Its Effect on the Visual Cliff Behavior of 1-Year-Olds. *Developmental Psychology*, *21*, 195–200.

Symons, L., Hains, S., Development, D. M.-I. B. and, & 1998, undefined. (n.d.). Look at me: Five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Elsevier*.

Tacon, P., & Pardoe, C. (2002). Dogs make us human. *Australian Museum*, *27*(4), 52–61.

Thalmann, O., Shapiro, B., & Cui, P. (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science*, *342*(6160), 871–874.

THEBERGE, J. B., & FALLS, J. B. (1967). Howling as a Means of Communication in Timber Wolves. *American Zoologist*, *7*(2), 331–338. https://doi.org/10.1093/icb/7.2.331

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-Directed Speech Facilitates Word Segmentation. *Infancy*, *7*(1), 53–71. https://doi.org/10.1207/s15327078in0701_5

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The Infant's Response to Entrapment between Contradictory Messages in Face-to-Face Interaction. *Pediatrics*, *62*(3).

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*(1), 2–7. https://doi.org/10.1016/j.specom.2006.10.003

Voegelin, C. F., & Robinett, F. M. (1954). "Mother Language" in Hidatsa. *International Journal of American Linguistics*, *20*(1), 65–70. https://doi.org/10.1086/464252

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, *9*(12), 585–594. https://doi.org/10.1016/J.TICS.2005.10.011

Waller, B. M., Warmelink, L., Liebal, K., Micheletta, J., & Slocombe, K. E. (2013). Pseudoreplication: A widespread problem in primate communication research. *Animal Behaviour*, *86*(2), 483–488. https://doi.org/10.1016/j.anbehav.2013.05.038

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue Canadienne de Psychologie*, *43*(2), 230–246. https://doi.org/10.1037/h0084224

Wilkins, A. S., Wrangham, R. W., & Tecumseh Fitch, W. (2014). The “domestication syndrome” in mammals: A unified explanation based on neural crest cell behavior and genetics. *Genetics*, *197*(3), 795–808. https://doi.org/10.1534/genetics.114.165423

Wrangham, R. W., & Glowacki, L. (2012). Intergroup Aggression in Chimpanzees and War in Nomadic Hunter-Gatherers: Evaluating the Chimpanzee Model. *Human Nature*. https://doi.org/10.1007/s12110-012-9132-1

Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013). Vowel Hyperarticulation in Parrot-, Dog- and Infant-Directed Speech. *Anthrozoos: A Multidisciplinary Journal of The Interactions of People & Animals*, *26*(3), 373–380. https://doi.org/10.2752/175303713X13697429463592

Yin, S. (2002). A new perspective on barking in dogs (Canis familaris.). *Journal of Comparative Psychology*, *116*(2), 189–193. https://doi.org/10.1037/0735-7036.116.2.189

Yin, S., & McCowan, B. (2004). Barking in domestic dogs: context specificity and individual identification. *Animal Behaviour*, *68*(2), 343–355. https://doi.org/10.1016/j.anbehav.2003.07.016

Young, L. J., & Wang, Z. (2004). The neurobiology of pair bonding. *Nature Neuroscience*, *7*(10), 1048–1054. https://doi.org/10.1038/nn1327