



OPEN

Multilevel fMRI adaptation for spoken word processing in the awake dog brain

Anna Gábor^{1,2✉}, Márta Gácsi^{2,3}, Dóra Szabó², Ádám Miklósi^{2,3}, Enikő Kubinyi² & Attila Andics^{1,2}

Human brains process lexical meaning separately from emotional prosody of speech at higher levels of the processing hierarchy. Recently we demonstrated that dog brains can also dissociate lexical and emotional prosodic information in human spoken words. To better understand the neural dynamics of lexical processing in the dog brain, here we used an event-related design, optimized for fMRI adaptation analyses on multiple time scales. We investigated repetition effects in dogs' neural (BOLD) responses to lexically marked (praise) words and to lexically unmarked (neutral) words, in praising and neutral prosody. We identified temporally and anatomically distinct adaptation patterns. In a subcortical auditory region, we found both short- and long-term fMRI adaptation for emotional prosody, but not for lexical markedness. In multiple cortical auditory regions, we found long-term fMRI adaptation for lexically marked compared to unmarked words. This lexical adaptation showed right-hemisphere bias and was age-modulated in a near-primary auditory region and was independent of prosody in a secondary auditory region. Word representations in dogs' auditory cortex thus contain more than just the emotional prosody they are typically associated with. These findings demonstrate multilevel fMRI adaptation effects in the dog brain and are consistent with a hierarchical account of spoken word processing.

During spoken word processing, the human brain separates lexical meaning from emotional prosody^{1–3}. Lexical processing entails speech sound sequence recognition and the matching of such sequences to previously associated meanings. This requires access to pre-existing speech sound sequence representations, assumedly involving higher levels of the speech processing hierarchy^{3,4}. In contrast, emotional prosody processing is largely based upon simple acoustic cues (such as pitch and pitch change)^{5–8}. In an fMRI study with awake dogs (*Canis familiaris*) listening to words, we found evidence that the ability to separately process lexical information and emotional prosody is not specific to humans⁹. Dogs showed an overall right hemispheric bias for lexically marked (praise) but not for lexically unmarked (neutral) words, independently of emotional prosody. While this initial study identified a set of auditory brain regions in dogs that are responsive to human speech in general, the distribution of labour among these regions remained unclear. To functionally characterize speech-responsive regions and better understand the relationship of lexical and emotional prosody processing in dogs, here we followed up directly on our previous work, using the same stimuli, but applying a multilevel fMRI adaptation paradigm.

Habituation/dishabituation paradigms are successfully used in various species, including dogs, to examine whether individuals are able to distinguish among certain stimuli^{10,11}. This behavioural priming phenomenon is often linked to a reduction in neural activity associated with repeated stimulus processing, which can be measured by single-cell recording¹², electrophysiological measures¹³ or haemodynamic imaging techniques like PET and fMRI¹⁴. fMRI adaptation effects (reduction in the BOLD signal after repeated presentations of a stimulus) have been demonstrated in different mammal species (e.g. macaques¹⁵, rats¹⁶, and also in humans¹⁷). fMRI adaptation can occur at different time scales from seconds¹⁸ to minutes^{19,20}. Short-term and long-term fMRI adaptation effects appear to be induced by different underlying mechanisms: short-term or rapid fMRI adaptation²¹ reflects stimulus similarity from the directly preceding stimuli, also referred to as carry-over effects²²,

¹MTA-ELTE 'Lendület' Neuroethology of Communication Research Group, Hungarian Academy of Sciences – Eötvös Loránd University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary. ²Department of Ethology, Eötvös Loránd University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary. ³MTA-ELTE Comparative Ethology Research Group, Hungarian Academy of Sciences – Eötvös Loránd University, Pázmány Péter sétány 1/C, 1117 Budapest, Hungary. ✉email: annagabor33@gmail.com

while long-term fMRI adaptation is thought to have a role in the formation of long-term stored representations and to thus reflect long-lasting neural sharpening for learned stimuli^{19,23}. Short-term repetition suppression has also been suggested to reflect initial responses²⁴, early sensory, mostly bottom-up processes, while long-term repetition suppression may reflect top-down modulation from regions higher in the processing hierarchy^{25,26}. Aging can modulate adaptation effects both neurally and behaviourally. The nature of such modulatory effects is, however, unclear. Age-related differences in fMRI adaptation in humans may be related to the reduction of neural selectivity in older individuals (i.e. neural dedifferentiation²⁷), or a decline in inhibitory processes that may result in inefficient filtering of irrelevant stimulus variation²⁸.

Auditory fMRI adaptation studies in humans suggest that lexical processing, typically tested by repeated presentations of known words, can be reflected by both long-term^{29–31} and short-term^{32,33} repetition suppression effects. Long-term priming for lexical processing has also been demonstrated behaviourally^{30,34,35}. Several areas of the human auditory cortex (e.g. BA 21³⁶) and the inferior frontal cortex (e.g. BA 45, 47³⁶) are more strongly adapted during lexical than during phonetic tasks, especially in the left hemisphere^{30,31,36}. Lexical meaning processing mostly occurs in temporal and frontal areas^{7,8,37}. Most human studies on lexical processing reported a clear left hemispheric bias, typically linked to higher levels (mid and anterior STG) of the auditory ventral stream^{3,4}.

Emotional prosody processing is highly dependent on acoustic features³⁷. After subcortical auditory regions provide a first acoustic analysis of vocal emotions, further integration and cognitive appraisal of the acoustic cues take place in the primary and secondary auditory cortices^{6,38,39}. The involvement of both subcortical and cortical auditory regions in processing human emotional vocalizations has also been demonstrated in dogs⁴⁰. Based on both human³⁷ and animal experiments^{41,42}, acoustic processing is reflected by adaptation effects already at an early stage of processing, in the subcortical auditory thalamus. In humans, acoustic sensitivity is often shown to be reflected by short-term adaptation effects for consecutive stimuli⁴³, however, long-term acoustic adaptation effects over several minutes have also been demonstrated^{123,44}.

To dogs, communicating effectively with humans and associating meanings to words is highly relevant^{45,46}, but very little is known about the similarities and differences between the auditory mechanisms involved in lexical processing in dogs and humans. Beyond our previous study on dogs' lexical processing⁹, there have been two recent dog fMRI studies that used words as stimuli, but neither of these two was designed to reveal lexical effects. One study found an increased activity for novel pseudowords compared to trained words in the broadly defined parietotemporal cortex, but that effect was related to novelty processing rather than to lexical processing⁴⁷. The other study showed that stimulus-reward neural associations are formed less effectively for verbal than for visual or olfactory cues⁴⁸. Although fMRI adaptation appears to provide an efficient means to investigate auditory processing mechanisms in a passive listening paradigm, it has never been exploited in dogs before.

In this fMRI experiment, dogs listened to lexically and prosodically marked and unmarked words in all combinations. This way we could separately investigate the effects of lexical and prosodic processing. The term lexically marked (meaningful) word refers to sound sequences that are typically used in the same context: when praising the dog. Lexically unmarked (meaningless) words are not associated with any specific contexts for dogs. We use the term lexical meaning to differentiate it from the intonationally conveyed meaning of a sound sequence—the latter one is reflected in emotional prosody. To avoid speaker-related familiarity and context difference effects, which strongly affect dogs' behaviour in responses to verbal commands⁴⁹, all words were spoken by a single female trainer, who often talked to all dogs during the several month-long fMRI training process. We used a rapid event-related design, presenting each word as a separate trial and modelling long-term (across the entire run, i.e. 30 repetitions over ~6.5 min) and short-term (across consecutive trials, i.e. 3 repetitions within 9 s) repetition effects, to measure fMRI adaptation at different time scales, similarly to previous works^{19,23,50}. We hypothesized that in dogs, similarly to humans, lexical and prosodic processing are reflected by distinct fMRI adaptation effects in speech-responsive auditory brain regions, and are modulated by age. More specifically, we predicted that in dogs lexical meaning-based adaptation (1) would be independent of prosody effects at higher levels of the processing hierarchy, and (2) would exhibit right hemisphere dominance.

Results

We found no significant effects (voxel-level, FWE-corrected $P < 0.05$) of either lexical meaning or emotional prosody with the classic 4-condition-based model, neither in whole-brain tests nor in functionally defined speech-responsive regions (Fig. 1).

Next, we performed short- and long-term fMRI adaptation analyses (see Methods for details). Significant main effects and interactions from these analyses are summarized in Table 1.

The prosody-based short-term fMRI adaptation analyses revealed a bilateral repetition effect in the tectum mesencephali (TM). Follow-up pairwise comparisons in the TM indicated a significant suppression effect between the second and third repetitions ($T_{11} = 3.907$, $P = 0.003$), but no difference between the first and second repetitions ($T_{11} = -1.638$, $P = 0.130$) (Fig. 2A). There were no significant repetition or hemisphere effects in any auditory cortical regions. The lexical meaning-based short-term fMRI adaptation analyses revealed no significant effects of either repetition or hemisphere, neither in subcortical nor in cortical speech-responsive regions.

Long-term fMRI adaptation analyses revealed prosody-dependent repetition effects in the subcortical TM and the near-primary auditory cortical mid ectosylvian sulcus (mESS); and lexical meaning-dependent repetition effects in the near-primary mESS and mid suprasylvian sulcus (mSSS), and the secondary auditory cortical caudal ectosylvian gyrus (cESG). Figure 2B displays long-term fMRI adaptation effects. Bar graphs display adaptation coefficient values: the larger the adaptation coefficient, the greater the repetition suppression effect. More specifically, in the subcortical TM, we found a repetition by prosody by hemisphere interaction, indicating repetition enhancement for praising prosody and repetition suppression for neutral prosody, with a larger

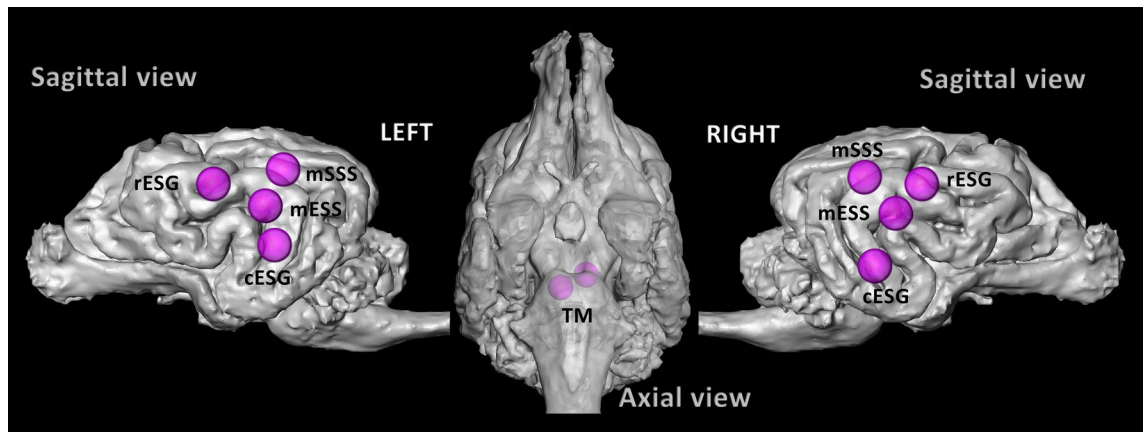


Figure 1. Speech-responsive auditory regions in the dog brain. Purple spheres ($R=4$ mm) are centred around previously functionally defined auditory activity peaks (Andics et al., 2016), using a speech vs. silence contrast at the group level with the same dog participants, and used as regional search spaces. Speech-responsive peaks were defined individually within the above spheres (see Supplementary Materials, Table S1), and individual ROIs with 2-mm-radius—used in the later analyses—were created around them. *TM* left tectum mesencephali, *mESS* mid ectosylvian sulcus, *mSSS* mid suprasylvian sulcus, *rESG* rostral ectosylvian gyrus, *cESG* caudal ectosylvian gyrus.

Brain region	Effect	df1	df2	F	P
Prosody-based short-term fMRI adaptation					
TM	Repetition	2	22	6.368	0.007
Lexical meaning-based short-term fMRI adaptation					
	–				
Long-term FMRI (fMRI) adaptation					
TM	Repetition \times prosody \times hemisphere	29	290	2.283	<0.001
mess (mESS)	Repetition \times lexical meaning \times hemisphere	29	290	1.859	0.006
	Repetition \times lexical meaning \times prosody	29	290	1.863	0.006
	Repetition \times lexical meaning \times Age (age)	29	290	2.030	0.002
mSSS	Repetition \times lexical meaning \times hemisphere	29	290	1.836	0.007
cESG	Repetition \times lexical meaning	29	290	2.365	<0.001

Table 1. FMRI adaptation effects for speech processing in dog auditory regions. The table lists significant effects of the prosody- and lexical meaning-based short-term fMRI adaptation analyses (RM ANOVAs with factors repetition and hemisphere), and the long-term fMRI adaptation analyses (RM ANOVAs with factors repetition, lexical meaning, prosody, and hemisphere; and covariate age). All analyses were performed in five speech-responsive regions (TM, mSSS, mESS, rESG, cESG), bilaterally (see Fig. 1). No suprathreshold main effects were found. In case of suprathreshold 3-way-interactions, 2-way-interactions with the same factors are not reported (but see text for details on post-hoc tests). Only effects surviving Bonferroni correction for multiple comparisons are reported. $N=12$.

difference between the two in the left hemisphere. Post-hoc tests for repetition effects in the TM, however, did not reach significance for either praising or neutral prosody, in either hemisphere ($F_s < 1.165$, $P_s > 0.260$). In mESS, we found multiple three-way interaction effects, involving the factors repetition, lexical meaning and, as a third factor, hemisphere, prosody or age. Post-hoc tests in the mESS revealed significant repetition suppression for praise words (1) in the right hemisphere ($F_{29,319} = 1.954$, $P = 0.003$); (2) in praising prosody ($F_{29,319} = 1.779$, $P = 0.010$), and (3) in younger dogs ($F_{29,145} = 2.78$, $P < 0.001$). Note that for the post-hoc test age was added as a category variable (young: 2–5 years, $N=6$; old: 7–10 years, $N=6$, Fig. S1). In mSSS, we found a repetition by lexical meaning by hemisphere interaction, suggesting that lexical meaning-dependent adaptation in this region was stronger for praise words, in the right hemisphere. Post-hoc tests of repetition effects in the mSSS, however, did not reach significance for either praise or neutral words, in either hemisphere ($F_s < 1.466$, $P_s > 0.061$). In cESG, we found a repetition by lexical meaning interaction. Post-hoc tests in the cESG revealed significant repetition suppression for praise words ($F_{29,319} = 2.179$, $P = 0.001$). No significant main effects or interactions were revealed in the rostral ectosylvian gyrus (rESG).

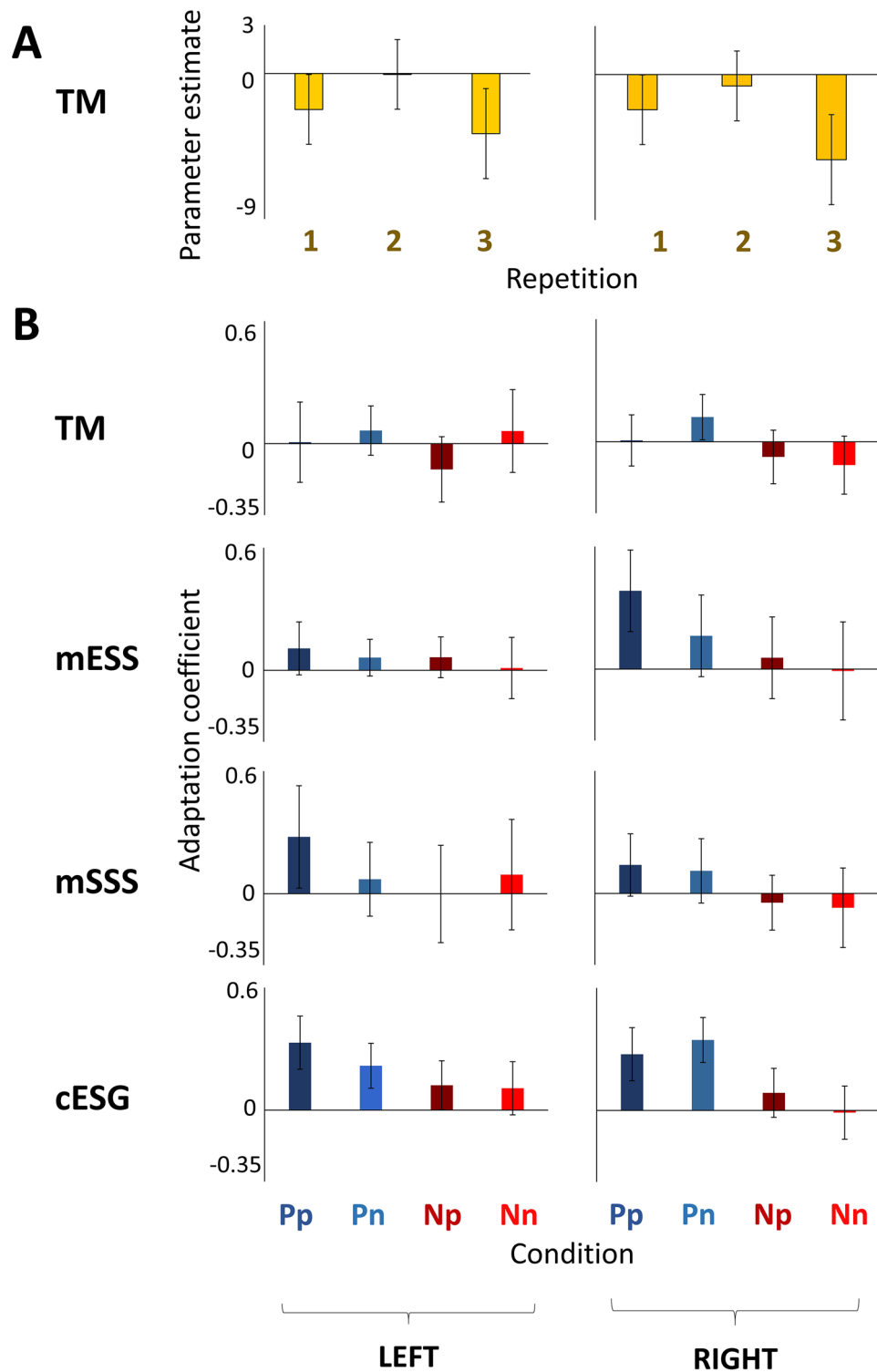


Figure 2. fMRI adaptation effects for speech processing in dogs. **(A)** Prosody-based short-term fMRI adaptation effects. Parameter estimates (trial-based beta values) are averaged for all trials that were the first, second, or third consecutive repetitions of the same prosody. **(B)** Long-term fMRI adaptation effects for prosody and lexical meaning. Adaptation coefficients are defined as the negative of the slope of the linear trendline for trial-based beta values across repetitions (see Methods for details). *Pp* lexically marked (praise) words with praising prosody, *Pn* lexically marked (praise) words with neutral prosody, *Np* lexically unmarked (neutral) words with praising prosody, *Nn* lexically unmarked (neutral) words with neutral prosody. * $P < 0.005$; ** $P < 0.001$. Error bars represent SEM. $N = 12$.

Discussion

This study presents the first demonstration of fMRI adaptation effects in the dog brain (but note a recent report of repetition enhancement effects⁵¹). We used these effects successfully to demonstrate the involvement of certain auditory regions in lexical and prosodic processing. By characterizing neural responses in dogs' speech-responsive brain regions for auditorily processed words using a multilevel fMRI adaptation paradigm, we found (1) lexical meaning-dependent long-term fMRI adaptation effects in near-primary and secondary auditory cortical regions, and (2) emotional prosody-dependent fMRI adaptation in a subcortical and a near-primary auditory cortical region. Lexical adaptation appeared only cortically and only as a long-term effect. Subcortical auditory regions showed only prosodic but no lexical adaptation. In a near-primary auditory cortical region lexical adaptation showed right-hemisphere bias, was enhanced by emotional prosody, and was modulated by age.

By analysing repetition effects, we demonstrated that three cortical speech-responsive auditory regions (mSSS, mESS, cESG) in dogs are sensitive to the lexical markedness of spoken words: these regions exhibited greater long-term fMRI adaptation—and thus a weaker overall response—for lexically marked (praise) words than for lexically unmarked (neutral) ones. In their studies on humans, Gagnepain et al. (2008)³⁰ and Orfanidou et al. (2006)²⁹ also found long-term fMRI adaptation (and long-term behavioural priming) for meaningful words in multiple areas of the non-primary auditory cortex (e.g. mSTG, pSTG, L MTG). Additionally, Gold et al. (2005)³⁶ found stronger long-term fMRI adaptation during a lexical meaning-related task than during a phonological task in the left middle temporal gyrus (L MTG) and the left inferior frontal gyrus (L IFG). In addition, we found no short-term lexical adaptation effects. Previous literature is inconclusive on whether in humans lexical processing is reflected also in short-term³³, or only in long-term adaptation effects^{29,30}. Nevertheless, short-term repetition effects are usually reported for the repetitions of simple, stimulus-dependent cues (such as emotional prosody cues), rather than for abstract stimulus properties (such as lexical meaning)^{19,24–26}. Our findings thus suggest that in dogs, as in humans^{29,30,52}, lexical processing is reflected in fMRI adaptation effects in a longer time scale in higher-level cortical regions.

The right hemisphere bias for lexical adaptation in the near-primary auditory cortex corroborates our earlier results⁹, suggesting that the processing of lexically marked words in dogs is more pronounced in the right hemisphere. Across auditory cortical regions, we found hemispheric bias only for lexical but not for prosodic markedness. In humans, lexical meaning processing shows hemispheric asymmetry towards the left hemisphere of the brain^{2,4}, while highly emotional speech stimuli are processed with lateral symmetry⁵ or with a right bias^{5,53}. In dogs, behavioural measures were used in two recent studies to search for the possible presence of functional hemispheric asymmetries for processing human speech. While no consistent head-turn bias was found for naturally spoken meaningful instruction words in either study^{54,55}, right head-turn bias (possibly indicating left hemispheric bias) was found for commands where meaningful phonemic cues were made salient artificially, and left head-turn bias (perhaps indicating a right hemispheric bias) for commands where emotional prosodic or speaker-related cues were made salient artificially⁵⁴. While the present fMRI findings corroborate earlier neuroimaging results⁹, it is harder to reconcile them with behavioural reports. One possibility is that behavioural measures of lateralization do not reflect functional hemispheric asymmetries as directly as it has often been proposed. A combined behavioural-fMRI lateralization study in humans also demonstrated that orienting biases for speech stimuli are not necessarily coupled with lateralized processing⁵⁶. Combined behavioural-fMRI investigations would have the capacity to reveal the neural pattern behind orienting biases. Another possible explanation for the seemingly contradicting findings is that the right bias for meaningful words presented here reflected the recognition of the processed lexical item (i.e. access to learned speech sound sequences), while the right head-turn bias in the behavioural study⁵⁴ may have revealed a left bias for segmental analysis (i.e. identifying phonemes in a speech stream), a necessary prerequisite of lexical processing. Left bias for segmental and right bias for suprasegmental processing is consistent with an acoustic account of lateralization (i.e. short vs long temporal windows for processing in left vs right auditory cortex, respectively)^{37,57}. We suggest that this account can explain many of the findings of Ratcliffe and Reby's (2014)⁵⁴ study. In our study, neither segmental nor speaker-related suprasegmental cues have been varied systematically, and prosodic suprasegmental cues did not lead to a hemispheric bias in the auditory cortex (and in the subcortical TM, the only region with hemisphere bias for prosodic adaptation, we found that prosodically more salient stimuli did not elicit stronger adaptation than neutral prosody in either hemisphere), so our findings neither support nor contradict the assumptions of the acoustic account of lateralization. Instead, our findings support a functional, meaningfulness-based account of lateralization. Hemispheric effects for processing meaningful, relevant sounds have been found in many species, including birds, non-primate mammals, and primates^{58–61}, even though most of these showed a left bias (but see⁶²). Note however that most of these studies tested conspecific sounds. It is possible that the recognition of a learned auditory stimulus elicits hemispheric bias in dogs, and while this bias is typically left-sided for conspecific vocal sounds, it becomes right-sided for vocalizations that elicit intense emotions (cf.^{55,60,63}).

Next to lexical adaptation, this study also showed evidence for emotional prosodic adaptation effects in the dog brain. We found short-term and long-term prosodic adaptation effects in a subcortical auditory region (TM), and long-term prosodic adaptation effect in a near-primary cortical auditory region (mESS). The involvement of the subcortical TM reflects the role of these early-stage areas in processing acoustic cues relevant to emotional prosody. According to single-unit experiments, in many species, the subcortical auditory thalamus shows short-term adaptation to stimulus repetitions^{41,42}. Anatomically, the speech-responsive TM region we used here involves the dog auditory thalamus, but the spatial resolution of the present study does not allow for its disentanglement from other, neighbouring subcortical structures. FMRI evidence for the involvement of early subcortical levels of the auditory pathway for processing vocal sounds has been reported for both humans³⁷ and dogs⁴⁰. The other prosody-sensitive region, mESS is centred around the sulcus located at the border of the mid ectosylvian gyrus (mESG), the primary auditory cortex of the dog, a region that receives tonotopic input from

the auditory thalamus⁶⁴. We found that mESS is the single cortical speech-responsive region where long-term fMRI adaptation was dependent not only on lexical meaning but also on prosody (being strongest for praise words in praising prosody). These findings suggest that the analysis of emotional prosody information in speech involves early levels of the auditory processing hierarchy in dogs.

Our findings suggest that dogs, similarly to humans, process emotional prosodic cues in spoken words at lower levels (subcortical and near-primary cortical regions, reflected in both short-term and long-term adaptation effects) and lexical information at higher levels (near-primary and secondary auditory cortical regions, reflected in long-term adaptation effects) of the auditory processing hierarchy. Prosody processing was thus subcortically independent of lexical cues, prosody influenced lexical processing in a near-primary cortical region and, finally, lexical processing was independent of prosodic cues in a secondary auditory cortical region. This hierarchical organization may reflect similarities of dog and human speech processing, but this does not imply that this processing hierarchy is of linguistic nature. Indeed, the prosodic-lexical hierarchy reported here and also in humans may reflect a more general, not speech-specific processing principle. According to Pessoa and Adolphs (2010)⁶⁵, perceptually salient (e.g. emotionally loaded, motivationally important) cues are typically analysed at lower levels (“low road”), and more complex, learnt, perceptually less salient cues of the same signal are analysed at higher levels (“high road”). This low road / high road processing hierarchy has been demonstrated in multiple species, independently of a linguistic context^{66–68}. In the present study, the prosodic manipulation was acoustically salient, as praising prosody was characterized by a higher pitch and pitch range than neutral prosody. In contrast, lexically marked and unmarked stimuli did not systematically differ in acoustic cues (contrasted conditions were matched for consonant–vowel structure and for emotional prosody), this learnt distinction was not salient acoustically. Note that there were also no familiarity differences between lexically marked and unmarked words because the lexically unmarked (neutral) words we selected here were words that had been used with a similar frequency to praise words in everyday speech, so the actual sound sequences were similarly familiar to dogs. Therefore, the only systematic difference between praise words and neutral words was that praise words were arbitrary sound sequences with an associated meaning, while neutral words were arbitrary sound sequences with no associated meaning. This contrast between emotional prosodic and lexical cues is not specific to our study—instead, it is a basic, essential difference between prosodic and lexical information and also applies to speech processing in humans.

So, does the reported lexical effect constitute evidence for human-analogue lexical representations in the dog brain? We do not suggest that the neural speech processing hierarchy shown here reflects any linguistic capacity in dogs. In contrast, our findings indicate that some of the neural mechanisms that support lexical processing may not be specific to humans. The reported lexical effect in dogs reveals differential processing of meaningful and meaningless words. Importantly, the fact that the presence of an associated meaning made a difference to the processing of a sound sequence in dogs, does not reveal lexical access. In other words, we do not know whether dogs learnt the lexical meaning (i.e. praise) associated to certain sound sequences (i.e. the praise words), or this association simply made those sound sequences more relevant to them and therefore easier to learn (and then recognize as known sound sequences). Consequently, we do not propose that dogs have human-analogue lexical representations, or that the observed lexical effects reveal complex or abstract processes. The only level of abstraction we argue for is regarding acoustics: although usually dogs only hear praise words in praising prosody, the lexical effect for praise words in secondary auditory regions was not stronger for praising than neutral prosody, suggesting that word representations in dogs’ auditory cortex thus contain more than just the emotional prosody they are typically associated with. The neural mechanisms underlying the reported lexical effects may not involve both core components of lexical processing (i.e. sound sequence recognition and meaning extraction). Based on this study alone we cannot claim that praise words, unlike neutral words, have been meaningful to dogs, it is also possible that the corresponding sound sequences were simply better learnt. Future studies will need to determine whether the neural process underlying this lexical effect reflects sound sequence recognition or word meaning extraction. Both accounts are plausible. Comparative behavioural work demonstrated that the dissociation of sound sequence and pitch during auditory processing is not unique to humans (e.g. dolphins⁶⁹, songbirds⁷⁰). But there is also evidence that at least some dogs can associate meaning to words (see^{71,72} for case studies with dogs correctly identifying hundreds of toys based on their name). Either way, the effect we report here evidences learning about words in dogs and cannot be accounted for by differences in acoustics or frequency-based familiarity.

We found no lexical or prosodic effects in a standard GLM-based analysis: this test revealed no brain region in dogs in which praise words or praising prosody elicited stronger or weaker overall activity than neutral words or prosody. This negative finding is not surprising, as the event-related design applied here, while more suitable to investigate cross-trial dynamics of brain responses, is known to be less robust to overall condition differences⁷³. Furthermore, the same contrasts did not show strong effects in speech-responsive auditory regions in a previous, block-design study either⁹. This shows that direct comparisons do not always constitute the optimal analysis of condition differences in fMRI: in the present case, an adaptation analysis was more informative.

The present study also showed that age modulates lexical adaptation effects in dogs. Specifically, the mESS adaptation effect difference between known and unknown words was larger in younger dogs. The small sample size of the present study does not allow for any strong conclusions on between-subject factors such as age. Nevertheless, it is worth noting that in humans, age effects have also been more pronounced in abstract, lexical/semantic components of repetition priming for language processing than in primary, perceptual components⁷⁴. Also, the reduced fMRI adaptation difference between conditions in older individuals supports the account that neural specificity decreases with age^{27,75}.

One limitation of the present study is that all stimuli were recorded from a single speaker, a female trainer of all tested dogs. While this might make our results less generalizable, we decided on using a single speaker with consideration to the reports that dogs process human vocal sounds in a highly context-sensitive manner⁷⁶, and that speaker familiarity affects their behavioural responses to instruction words⁴⁹. We aimed at using identical

stimuli across participants and also maximizing the relevance of our stimuli in this sense, similarly to other studies using a trainer's voice instead of a set of less familiar speakers (cf.⁷¹). Human fMRI studies on speech processing also often use a single speaker^{77,78}. One could argue that overall adaptation effects for speech stimuli may be different for unfamiliar speakers⁷⁹. Crucially, however, the adaptation effects here were all condition-dependent, that is, stimuli in one condition elicited stronger adaptation than stimuli in another condition, even though all were spoken by the same speaker. We cannot draw conclusions about the across-speaker generalizability of lexical representations in dogs based on the present study, but this does not question the lexical nature of the lexical adaptation effects we demonstrated here. Another limitation is that the lexically marked words we chose were all used in a single context: when rewarding the dog. Therefore, based on this study alone we cannot determine whether the revealed right bias (and other lexical effects) reflects lexical processing (sound sequence recognition or meaning extraction) in general or, more specifically, the processing of praise words spoken by a familiar person. To better understand the mechanisms of spoken word processing in dogs, further fMRI studies are required. In these future studies it will be important to test the role of speaker familiarity and the processing of words learned in different contexts (e.g. applying object names or instructions). Widening the framework in which dogs' neural responses to human vocal/verbal communication is investigated would be intriguing because the dog has been recently suggested as a complementary model species to the traditionally used primate and rodent models due to its evolutionary and ontogenetic development in the human social environment and also for many practical reasons (e.g., non-invasive measurements, ethical issues)^{46,80–82}.

This study demonstrated the usefulness of a multilevel fMRI adaptation approach to functionally characterize speech-responsive regions in the dog brain. We identified speech-responsive auditory regions involved in lexical and emotional prosody processing in dogs. We replicated our earlier findings⁹ that in dogs, lexically marked praise words are processed with a right-hemisphere bias. Lexical and prosodic adaptation patterns differed both temporally (long-term effects mostly for lexical processing and short-term effects only for emotional prosody processing) and spatially (lexical processing only cortically, in near-primary and secondary speech-responsive auditory regions, and emotional prosody processing only in subcortical and near-primary auditory cortical regions), suggesting that they indeed reflected distinct stages of an auditory processing hierarchy in the dog brain. Our findings thus provide evidence for the hierarchical processing of spoken words in a speechless species.

Materials and methods

Participants. We tested 12 pet dogs (mean age (year) \pm SD 6.17 ± 2.82 , range 2–10 years; 3 breeds: 6 border collies, 5 golden retrievers, 1 German shepherd; 8 males and 4 females) living in human families. As this work is a follow-up on a previous study⁹, we used the same dog participants (all but one dog participated in the previous study; two dogs that participated in the previous study were not available for measurements any more). The training procedure for dogs to lie motionless throughout the test was based on individual and social learning using positive reinforcement and has been described in detail previously⁴⁰.

Stimuli. The stimuli were lexically marked (praise) words, meaningful for the dogs, and lexically unmarked (neutral) words, meaningless for the dogs, with praising and neutral prosody in all combinations, identical to those used in Andics et al. (2016)⁹. The three lexically marked (praise) words in Hungarian were: azaz ['ɒzɒz] / ügyes ['yɟɛʃ] / jól van ['jo:lvan] for "that's it / clever / well done", all used to praise the tested dogs. As lexically unmarked (neutral) words, we used three conjunction words: akár ['ɒkɒr] / olyan ['oɟɒn] / mégsem ['me:ʃsem] for "as if / such / yet", used with similar frequency in everyday speech, but not used in dog-directed speech. We recorded all six words, both with praise and neutral prosody twice (24 recordings in total). A female trainer of the dogs (MG) spoke the words, and she was always present at the scanner during the test sessions. The praising prosody stimuli were characterized by higher pitch and greater pitch range than the neutral prosody stimuli (praising / neutral prosody: mean (F0) = $268(\pm 20)/165(\pm 6)$ Hz, $F_{1,20} = 289.725$, $P < 0.001$; mean (F0 range) = $277(\pm 93)/46(\pm 9)$ Hz, $F_{1,20} = 68.264$, $P < 0.001$), independently of lexical markedness. There were no systematic pitch or pitch range differences between lexically marked (praise) words and lexically unmarked (neutral) words. To ensure that the stimulus voice has typical acoustic variation, we recorded the same praise and neutral words with both prosodies from 14 other persons and compared the pitch parameters of these reference voices across conditions via RM ANOVA. Here, again, words with praising prosody had higher pitch and higher pitch range than words with neutral prosody (praising / neutral prosody: mean (F0) = $216(\pm 67)/161(\pm 55)$ Hz, $F_{1,13} = 67.122$, $P < 0.001$; mean (F0 range) = $144(\pm 71)/37(\pm 18)$ Hz, $F_{1,13} = 44.032$, $P < 0.001$), but we found no systematic acoustic differences between praise and neutral words (all $F_s < 1$).

fMRI experimental design. We used four speech conditions (with three words per condition): (1) lexically marked (praise) words with praising prosody (Pp), (2) lexically marked (praise) words with neutral prosody (Pn), (3) lexically unmarked (neutral) words with praising prosody (Np), and (4) lexically unmarked (neutral) words with neutral prosody (Nn). We also added a silent condition and used it as a baseline in later analyses. A semi-continuous event-related fMRI paradigm was applied, in which each stimulus was played in 1 s long silent gaps (one stimulus per gap) between 2 s long volume acquisitions. Stimulus onsets were at 0.05 s within the silent gaps. Word lengths were between 0.484–0.896 s (0.642 s on average). One measurement consisted of 135 stimulus presentations (30 of each main condition, with 10 repetitions of every single word with both prosodic patterns, and 15 silent events). A semi-random stimulus order was used, with the proviso that two consecutive stimuli are not the same words with the same prosody. Conditions were evenly distributed, but the order was otherwise random and varied across participants. The experiment consisted of a single approximately 6.5-min run for each dog (the total duration of the run is limited by how long dogs can be instructed to lay motionless).

Stimulus order	Pp	Pn	Pp	Sil	Nn	Pn	Nn	Np	Sil	Np	Nn	...	Pp
Prosody-based STA:	p1	n1	p1	-	n1	n2	n3	p1	-	p1	n1	...	p1
Lexical meaning-based STA:	P1	P2	P3	-	N1	P1	N1	N2	-	N1	N2	...	P1
Condition-dependent LTA:	Pp1	Pn1	Pp2	-	Nn1	Pn2	Nn2	Np1	-	Np2	Nn3	...	Pp30

Table 2. Illustration of coding condition-dependent long-term repetitions, and prosody-based and lexical meaning-based short-term repetitions. *STA* short-term fMRI adaptation; *LTA* long-term fMRI adaptation; *P* lexically marked (praise) words; *N* lexically unmarked (neutral) words; *p* praising prosody; *n* neutral prosody; *Sil* silence.

Scanning procedure. During scanning, the stimulus presentation was controlled by MatLab (version 9.1) Psychophysics Toolbox 3⁸³ and synchronized with volume acquisitions by TTL trigger pulses. Stimuli were presented via MRI-compatible sound-attenuating headphones (MR Confon) that also protected the ears of the dogs from scanner noises. A Philips 3 T whole-body scanner and a Philips SENSE Flex Medium coil were used to perform the measurements, at the MR Research Centre of the Semmelweis University, Budapest. For functional scans, we used a single-shot gradient-echo planar imaging (EPI) sequence to acquire volumes of 29 transverse slices, with 0.5 mm gaps, covering the whole brain (slice order: ascending; spatial resolution including slice gaps: 3.5 × 3.5 × 3.5 mm; TR: 3.0 s; TE: 36 ms; flip angle: 90°; 64 × 64 matrix). One measurement consisted of 139 volumes. A T1-weighted anatomical brain image was acquired in a separate session (turbo-field echo (TFE) sequence; spatial resolution: 1 × 1 × 1 mm, 180 slices).

Our subjects had previously been trained to lie motionless for ~ 8 min without any restriction. We applied an absolute head motion threshold of 2 mm (for each translation direction) and 2° (for each rotation direction) across the entire run. To search for possible condition effects on head motion, we calculated framewise displacement (FD) in each dog (mean FD = 0.23 (± 0.15) mm)^{84–86}. This average FD value is comparable to a typical human adult's movement parameters measured in event-related task fMRI studies⁸⁷. Head motions following sound and silence conditions did not differ ($T_{14} = -0.836$, $p = 0.417$). RM ANOVA on dogs' FD values revealed no systematic differences in head motion across acoustic conditions (lexical meaning: $F_{1,29} = 0.009$, $p = 0.926$; prosody: $F_{1,29} = 2.129$, $p = 0.155$; lexical meaning × prosody: $F_{1,29} = 0.050$, $p = 0.825$).

fMRI data coding and statistical analysis. fMRI data preprocessing and analysis were performed using the SPM8 toolbox (www.fil.ion.ucl.ac.uk/spm) of MATLAB R2013a (<https://www.mathworks.com/products/matlab/>). Preprocessing procedure was identical to that in⁹ and involved manual and automatic spatial realignment, coregistration, normalization to an anatomical template, and smoothing. Individual statistical maps were obtained based on the general linear model. We specified two models. For a standard analysis with condition-based contrasts, we modelled the 5 main conditions (Pp, Pn, Np, Nn, and Sil) and used condition regressors. In a second GLM for the fMRI adaptation analyses, each trial (4 speech conditions × 30 repetitions) was modelled separately, and we used trial-based regressors.

To individuate the functional localization of speech-responsive auditory brain areas, we defined with the help of a previous study⁹ with the same stimuli and participants. There, group-level activity peaks of the speech (all conditions) vs silence contrast included the following bilateral auditory subcortical and cortical regions: left and right tectum mesencephali (L TM: - 4, -, - 12 R TM: 2, - 12, - 10), mid suprasylvian sulcus (L mSSS: - 16, - 14, 16; R mSSS: 18, - 14, 14), mid ectosylvian sulcus (L mESS: - 28, - 10, 8; R mESS: 22, - 6, 6), rostral ectosylvian gyrus (L rESG: - 22, 2, 14; R rESG: 20, - 2, 14), caudal ectosylvian gyrus (L cESG: - 24, - 10, - 2; R cESG: 26, - 10, - 6) (Fig. 1). These coordinates (mm) denote left to right, posterior to anterior, and inferior to superior directions respectively, using the same dog brain template space as in Andics et al. (2016)⁹. We created spheres ($r = 4$ mm) around these peaks and used them as regional search spaces. We then determined the speech-responsive peak within each of these regional search spaces at the individual level (for a list of individual coordinates, see Table S1). In case of a single dog (D12) who did not participate in the previous study, we used group-level peaks. We created spherical ROIs around these individually specified peaks ($r = 2$ mm). Therefore, each dog had a unique set of ROIs, which were nevertheless determined within the group analysis-based regional search spaces. A similar method was used in a recent dog fMRI study by our group, see⁵¹. Using trial-based regressors, we then determined parameter estimates (beta values) for each event, averaged within the above described, individual ROIs using WFU PickAtlas⁸⁸.

To investigate long-term condition-dependent fMRI adaptation, we coded each event with reference to the number of preceding repetitions (1 to 30) of the same condition within the test run (e.g. Pp1, Pp2, Pp3... Pp30) (Table 2). We then compared event-specific parameter estimates (beta values) using RM ANOVAs with hemisphere (left, right), lexical meaning (P, N), prosody (p, n) and repetition (1, 2 ... 30) as within-subject factors and age as a covariant within each speech-responsive region. We applied Bonferroni correction for the number of ANOVAs performed (i.e. the number of bilateral ROIs). If a significant interaction with repetition was found, we carried out follow-up tests to investigate repetition effects for each level of the other contributing factors. To illustrate long-term fMRI adaptation effects, adaptation coefficient values were calculated (Fig. 2B): first, we fitted a linear trendline to parameter estimates across repetitions, for each condition; second, we calculated the slope of this trendline, i.e., the rate of BOLD response decrease; third, we took the negative of this slope. Correspondingly, the larger the adaptation coefficient value, the greater the fMRI adaptation (repetition suppression) effect. In an additional analysis, we performed another series of RM ANOVAs to search for possible short-term fMRI adaptation effects for lexical meaning and prosody. For this, we separately coded trials for lexical meaning (P

included Pp and Pn, N included Np and Nn) and prosody (p included Pp and Np, n included Pn and Nn). To investigate short-term lexical meaning-based fMRI adaptation we coded every trial based on lexical meaning (P, N) and repetition, i.e. the number of directly preceding trials with the same lexical meaning (1, 2, 3). For example, P2 referred to a praise word that was the second consecutive repetition of the same lexical meaning. We included events until up to 3 repetitions, as 4 or more consecutive repetitions of the same lexical meaning were rare (2.5% of all cases).

To investigate short-term prosody-based fMRI adaptation we coded events similarly, but now based on prosody (p, n) and repetition (1, 2, 3) (Table 2). We then applied RM ANOVAs on beta values for each speech-responsive region with the factors hemisphere and repetition. Only effects that survive Bonferroni correction are reported.

Ethical statement. Research was done in accordance with the Hungarian regulations on animal experimentation and the Guidelines for the use of animals in research described by the Association for the Study Animal Behaviour (ASAB). Ethical approval was obtained from the local ethical committee (Állatkísérleti Tudományos Etikai Tanács KA-1719 / PEI/001/1,490–4/2015, Budapest, Hungary; Pest Megyei Kormányhivatal Élelmiszerlánc-Biztonsági és Állategészségügyi Igazgatósága XIV-I-001/520–4/2012, Budapest, Hungary). The dog owners were volunteers who received no monetary compensation and gave their written consent to participate with their dogs in the study.

Received: 8 January 2020; Accepted: 30 June 2020

Published online: 03 August 2020

References

- Shtyrov, Y., Pihko, E. & Pulvermüller, F. Determinants of dominance: Is language laterality explained by physical or linguistic features of speech?. *Neuroimage* **27**, 37–47 (2005).
- Devlin, J. T., Matthews, P. M. & Rushworth, M. F. S. Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *J. Cogn. Neurosci.* **15**, 71–84 (2003).
- DeWitt, I. & Rauschecker, J. P. Phoneme and word recognition in the auditory ventral stream. *Proc. Natl. Acad. Sci. USA* **109**, E505–E514 (2012).
- Binder, J. R., Desai, R. H., Graves, W. W. & Conant, L. L. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* **19**, 2767–2796 (2009).
- Frühholz, S., Trost, W. & Kotz, S. A. The sound of emotions—Towards a unifying neural network perspective of affective sound processing. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2016.05.002> (2016).
- Pannese, A., Grandjean, D. & Frühholz, S. Subcortical processing in auditory communication. *Hear. Res.* **328**, 67–77 (2015).
- de Heer, W. A., Huth, A. G., Griffiths, T. L., Gallant, J. L. & Theunissen, F. E. The hierarchical cortical organization of human speech processing. *J. Neurosci.* **37**, 6539–6557 (2017).
- Specht, K. Mapping a lateralization gradient within the ventral stream for auditory speech perception. *Front. Hum. Neurosci.* **7**, 629 (2013).
- Andics, A. *et al.* Neural mechanisms for lexical processing in dogs. *Scienc.* **353**, 1030–1032 (2016).
- Rankin, C. H. *et al.* Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.* <https://doi.org/10.1016/j.nlm.2008.09.012> (2009).
- Maros, K. *et al.* Dogs can discriminate barks from different situations. *Appl. Anim. Behav. Sci.* **114**, 159–167 (2008).
- Sobotka, S. & Ringo, J. L. Stimulus specific adaptation in excited but not in inhibited cells in inferotemporal cortex of Macaque. *Brain Res.* [https://doi.org/10.1016/0006-8993\(94\)90061-2](https://doi.org/10.1016/0006-8993(94)90061-2) (1994).
- Henson, R. N., Rylands, A., Ross, E., Vuilleumier, P. & Rugg, M. D. The effect of repetition lag on electrophysiological and haemodynamic correlates of visual object priming. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2003.12.020> (2004).
- Henson, R. N. A. & Rugg, M. D. Neural response suppression, haemodynamic repetition effects, and behavioural priming. *Neuropsychologica* **41**, 263–270 (2003).
- Kar, K. & Krekelberg, B. Testing the assumptions underlying fMRI adaptation using intracortical recordings in area MT. *Cortex* **80**, 21–34 (2016).
- Schafer, J. R., Kida, I., Rothman, D. L., Hyder, F. & Xu, F. Adaptation in the rodent olfactory bulb measured by fMRI. *Magn. Reson. Med.* <https://doi.org/10.1002/mrm.20588> (2005).
- Grill-Spector, K., Henson, R. & Martin, A. Repetition and the brain: Neural models of stimulus-specific effects. *Trends Cogn. Sci.* **10**, 14–23 (2006).
- Matsumoto, A., Iidaka, T., Haneda, K., Okada, T. & Sadato, N. Linking semantic priming effect in functional MRI and event-related potentials. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2004.09.008> (2005).
- Epstein, R. A., Parker, W. E. & Feiler, A. M. Two Kinds of fMRI Repetition Suppression? Evidence for Dissociable Neural Mechanisms. *J. Neurophysiol.* **99**, 2877–2886 (2008).
- Henson, R., Shallice, T. & Dolan, R. Neuroimaging evidence for dissociable forms of repetition priming. *Science* **287**, 1269–1272 (2000).
- Grill-Spector, K. & Malach, R. fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychol. (Amst.)* [https://doi.org/10.1016/S0001-6918\(01\)00019-1](https://doi.org/10.1016/S0001-6918(01)00019-1) (2001).
- Aguirre, G. K. Continuous carry-over designs for fMRI. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2007.02.005> (2007).
- Andics, A. *et al.* Neural mechanisms for voice recognition. *Neuroimage* **52**, 1528–1540 (2010).
- Sawamura, H., Orban, G. A. & Vogels, R. Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the fMRI adaptation paradigm. *Neuron* <https://doi.org/10.1016/j.neuron.2005.11.028> (2006).
- Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. B.* <https://doi.org/10.1098/rstb.2005.1622> (2005).
- James, T. W. & Gauthier, I. Repetition-induced changes in BOLD response reflect accumulation of neural activity. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.20165> (2006).
- Goh, J. O., Suzuki, A. & Park, D. C. Reduced neural selectivity increases fMRI adaptation with age during face discrimination. *Neuroimage* **51**, 336–344 (2010).
- Fabiani, M., Low, K. A., Wee, E., Sable, J. J. & Gratton, G. Reduced suppression or labile memory? Mechanisms of inefficient filtering of irrelevant information in older adults. *J. Cogn. Neurosci.* **18**, 637–650 (2006).

29. Orfanidou, E., Marslen-Wilson, W. D. & Davis, M. H. Neural response suppression predicts repetition priming of spoken words and pseudowords. *J. Cogn. Neurosci.* **18**, 1237–1252 (2006).
30. Gagnepain, P. *et al.* Spoken word memory traces within the human auditory cortex revealed by repetition priming and functional magnetic resonance imaging. *J. Neurosci.* **28**, 5281–5289 (2008).
31. Weber, K., Lau, E. F., Stillerman, B. & Kuperberg, G. R. The Yin and the Yang of prediction: An fMRI study of semantic predictive processing. *PLoS ONE* **11**, E148637 (2016).
32. Devauchelle, A. D., Oppenheim, C., Rizzi, L., Dehaene, S. & Pallier, C. Sentence syntax and content in the human temporal lobe: An fMRI adaptation study in auditory and visual modalities. *J. Cogn. Neurosci.* <https://doi.org/10.1162/jocn.2009.21070> (2009).
33. Kotz, S. A., Cappa, S. F., Von Cramon, D. Y. & Friederici, A. D. Modulation of the lexical-semantic network by auditory semantic priming: An event-related functional MRI study. *Neuroimage* <https://doi.org/10.1006/nimg.2002.1316> (2002).
34. Joordens, S. & Becker, S. The long and short of semantic priming effects in lexical decision. *J. Exp. Psychol. Learn. Mem. Cogn.* **23**, 1083–1105 (1997).
35. Becker, S., Moscovitch, M., Behrmann, M. & Joordens, S. Long-term semantic priming: a computational account and empirical evidence. *J. Exp. Psychol. Learn. Mem. Cogn.* **23**, 1059–1082 (1997).
36. Gold, B. T., Balota, D. A., Kirchoff, B. A. & Buckner, R. L. Common and dissociable activation patterns associated with controlled semantic and phonological processing: Evidence from fMRI adaptation. *Cereb. Cortex* <https://doi.org/10.1093/cercor/bhi024> (2005).
37. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
38. Moerel, M., De Martino, F., Ugurbil, K., Yacoub, E. & Formisano, E. Processing of frequency and location in human subcortical auditory structures. *Sci. Rep.* **5**, 17048 (2015).
39. Scott, S. K. & Wise, R. J. S. The functional neuroanatomy of prelexical processing in speech perception. *Cognition* **92**, 13–45 (2004).
40. Andics, A., Gácsi, M., Faragó, T., Kis, A. & Miklósi, Á. Voice-sensitive regions in the dog and human brain are revealed by comparative fMRI. *Curr. Biol.* **24**, 574–578 (2014).
41. Richardson, B. D., Hancock, K. E. & Caspary, D. M. Stimulus-specific adaptation in auditory thalamus of young and aged awake rats. *J. Neurophysiol.* **110**, 1892–1902 (2013).
42. Anderson, L. A., Christianson, G. B. & Linden, J. F. Stimulus-specific adaptation occurs in the auditory thalamus. *J. Neurosci.* **29**, 7359–7363 (2009).
43. Bestelmeyer, P. E. G., Latinus, M., Rouger, J., Maurice, P. & Belin, P. Adaptation to vocal expressions reveals multistep perception of auditory emotion. *J. Neurosci.* **34**, 8098–8105 (2014).
44. Polich, J. & McIsaac, H. K. Comparison of auditory P300 habituation from active and passive conditions. *Int. J. Psychophysiol.* [https://doi.org/10.1016/0167-8760\(94\)90052-3](https://doi.org/10.1016/0167-8760(94)90052-3) (1994).
45. Miklósi, A. *Dog Behaviour, Evolution, and Cognition* (Oxford University Press, Oxford, 2015).
46. Andics, A. & Miklósi, Á. Neural processes of vocal social perception: Dog-human comparative fMRI studies. *Neurosci. Biobehav. Rev.* **85**, 54–64 (2018).
47. Prichard, A., Cook, P. F., Spivak, M., Chhibber, R. & Berns, G. S. Awake fMRI reveals brain regions for novel word detection in dogs. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2018.00737> (2018).
48. Prichard, A., Chhibber, R., Athanassiades, K., Spivak, M. & Berns, G. S. Fast neural learning in dogs: a multimodal sensory fMRI study. *Sci. Rep.* <https://doi.org/10.1038/s41598-018-32990-2> (2018).
49. Kerepesi, A., Dóka, A. & Miklósi, Á. Dogs and their human companions: The effect of familiarity on dog-human interactions. *Behav. Processes* <https://doi.org/10.1016/j.beproc.2014.02.005> (2015).
50. Andics, A., McQueen, J. M. & Petersson, K. M. Mean-based neural coding of voices. *Neuroimage* **79**, 351–360 (2013).
51. Boros, M. *et al.* Repetition enhancement to voice identities in the dog brain. *Sci. Rep.* <https://doi.org/10.1038/s41598-020-60395-7> (2020).
52. Gold, B. T. & Buckner, R. L. Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. *Neuron* **35**, 803–812 (2002).
53. Wildgruber, D. *et al.* Identification of emotional intonation evaluated by fMRI. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2004.10.034> (2005).
54. Ratcliffe, V. F. & Reby, D. Orienting asymmetries in dogs' responses to different communicatory components of human speech. *Curr. Biol.* **24**, 2908–2912 (2014).
55. Reinholz-Trojan, A., Włodarczyk, E., Trojan, M., Kulczyński, A. & Stefańska, J. Hemispheric specialization in domestic dogs (*Canis familiaris*) for processing different types of acoustic stimuli. *Behav. Processes* <https://doi.org/10.1016/j.beproc.2012.07.001> (2012).
56. Fischer, J. *et al.* Orienting asymmetries and lateralized processing of sounds in humans. *BMC Neurosci.* **10**, 14 (2009).
57. Poeppel, D., Idsardi, W. J. & Van Wassenhove, V. Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. B* <https://doi.org/10.1098/rstb.2007.2160> (2008).
58. Denenberg, V. H. Hemispheric laterality in animals and the effects of early experience. *Behav. Brain Sci.* <https://doi.org/10.1017/S0140525X00007330> (1981).
59. Hopkins, W. D., Morris, R. D., Savage-Rumbaugh, E. S. & Rumbaugh, D. M. Hemispheric priming by meaningful and nonmeaningful symbols in language-trained chimpanzees (*Pan troglodytes*): further evidence of a left hemisphere advantage. *Behav. Neurosci.* <https://doi.org/10.1037/0735-7044.106.3.575> (1992).
60. Siniscalchi, M., Quaranta, A. & Rogers, L. J. Hemispheric specialization in dogs for processing different acoustic stimuli. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0003349> (2008).
61. Quaranta, A., Siniscalchi, M. & Vallortigara, G. Asymmetric tail-wagging responses by dogs to different emotive stimuli. *Curr. Biol.* <https://doi.org/10.1016/j.cub.2007.02.008> (2007).
62. Gil-da-Costa, R. & Hauser, M. D. Vervet monkeys and humans show brain asymmetries for processing conspecific vocalizations, but with opposite patterns of laterality. *Proc. R. Soc. B Biol. Sci.* **273**, 2313–2318 (2006).
63. Siniscalchi, M., d'Ingeo, S. & Quaranta, A. Lateralized Functions in the Dog Brain. *Symmetry* **9**, 71 (2017).
64. Evans, H. E. & deLahunta, A. *Miller's Anatomy of the Dog, 4th Edition. Miller's Anatomy of the Dog, Fourth Edition* (2013).
65. Pessoa, L. & Adolphs, R. Emotion processing and the amygdala: From a 'low road' to 'many roads' of evaluating biological significance. *Nat. Rev. Neurosci.* <https://doi.org/10.1038/nrn2920> (2010).
66. Kikuchi, Y., Horwitz, B. & Mishkin, M. Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. *J. Neurosci.* <https://doi.org/10.1523/jneurosci.2267-10.2010> (2010).
67. Carrasco, A. & Lomber, S. G. Evidence for hierarchical processing in cat auditory cortex: nonreciprocal influence of primary auditory cortex on the posterior auditory field. *J. Neurosci.* <https://doi.org/10.1523/jneurosci.2905-09.2009> (2009).
68. Rauschecker, J. P. & Scott, S. K. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nat. Neurosci.* <https://doi.org/10.1038/nn.2331> (2009).
69. Ralston, J. V. & Herman, L. M. Perception and Generalization of Frequency Contours by a Bottlenose Dolphin (*Tursiops truncatus*). *J. Comp. Psychol.* <https://doi.org/10.1037/0735-7036.109.3.268> (1995).
70. Sen, K., Theunissen, F. E. & Doupe, A. J. Feature Analysis of Natural Sounds in the Songbird Auditory Forebrain. *J. Neurophysiol.* **86**, 1445–1458 (2001).
71. Kaminski, J., Call, J. & Fischer, J. Word learning in a domestic dog: evidence for 'fast mapping'. *Science* **304**, 1682–1683 (2004).
72. Pillely, J. W. & Reid, A. K. Border collie comprehends object names as verbal referents. *Behav. Processes* **86**, 184–195 (2011).

73. Petersen, S. E. & Dubis, J. W. The mixed block/event-related design. *NeuroImage* <https://doi.org/10.1016/j.neuroimage.2011.09.084> (2012).
74. Rybash, J. M. Implicit memory and aging: a cognitive neuropsychological perspective. *Dev. Neuropsychol.* <https://doi.org/10.1080/87565649609540644> (1996).
75. Goh, J. O. S. Functional dedifferentiation and altered connectivity in older adults: neural accounts of cognitive aging. *Aging Dis.* **2**, 30–48 (2011).
76. Mills, D. S. What's in a word? A review of the attributes of a command affecting the performance of pet dogs. *Anthrozoos* <https://doi.org/10.2752/089279305785594108> (2005).
77. Binder, J. R. Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* **10**, 512–528 (2000).
78. Lawyer, L. & Corina, D. An investigation of place and voice features using fMRI-adaptation. *J. Neurolinguistics* <https://doi.org/10.1016/j.jneuroling.2013.07.001> (2014).
79. Latinus, M., Crabbe, F. & Belin, P. Learning-induced changes in the cerebral processing of voice identity. *Cereb. Cortex* <https://doi.org/10.1093/cercor/bhr077> (2011).
80. Bunford, N., Andics, A., Kis, A., Miklósi, Á. & Gácsi, M. *Canis familiaris* As a Model for Non-Invasive Comparative Neuroscience. *Trends Neurosci.* **40**, 438–452 (2017).
81. Bódizs, R., Kis, A., Gácsi, M. & Topál, J. Sleep in the dog: comparative, behavioral and translational relevance. *Curr. Opin. Behav. Sci.* <https://doi.org/10.1016/j.cobeha.2019.12.006> (2020).
82. Szabo, D. *et al.* Resting-state fMRI data of awake dogs (*Canis familiaris*) via group-level independent component analysis reveal multiple, spatially distributed resting-state networks. *bioRxiv* <https://doi.org/10.1101/409532> (2018).
83. Kleiner, M. *et al.* What's new in Psychtoolbox-3?. *Perception* **36**, S14 (2007).
84. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2011.10.018> (2012).
85. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2013.08.048> (2014).
86. Parkes, L., Fulcher, B., Yücel, M. & Fornito, A. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *Neuroimage* <https://doi.org/10.1016/j.neuroimage.2017.12.073> (2018).
87. Siegel, J. S. *et al.* Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. *Hum. Brain Mapp.* <https://doi.org/10.1002/hbm.22307> (2014).
88. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* **19**, 1233–1239 (2003).

Acknowledgements

This work was supported by the Hungarian Academy of Sciences [a grant to the MTA-ELTE Comparative Ethology Research Group (grant number F01/031); a grant to the MTA-ELTE 'Lendület' Neuroethology of Communication Research Group (grant number 95025) and a Bolyai Research Scholarship to AA and EK]; the Eötvös Loránd University; the Hungarian Scientific Research Fund (grant number LP2017-13/2017; the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant Number 680040); the National Research, Development, and Innovation Office (Grant NUMBER 115862K). AG and EK were supported through the National Excellence Program of the Ministry of Human Capacities (Grant Numbers ÚNKP-16-3: ELTE/8495/35(2016), ÚNKP-18-4, respectively). ÁM was supported through the Eötvös Loránd University Institutional Excellence Program (Grant Number 783-3/2018/FEKUTSRAT) of the Hungarian Ministry of Human Capacities and the National Brain Program (Grant Number NKFIH NPK_22-2022-0001). We thank Kálmán Czeibert for brain illustrations and the owners and their dogs for their participation.

Author contributions

A.G.: study conception and design, data collection, analysis and interpretation of data, writing of the manuscript. M.G.: study conception and design, data collection, dog training, interpretation of data, critical revision. D.S.: data collection, dog training, interpretation of data, critical revision. Á.M.: study conception and design, critical revision. E.K.: interpretation of data, critical revision. A.A.: study conception and design, data collection, interpretation of data, critical revision.

Competing interests

We declare that the authors have no competing interests as defined by Scientific Reports, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-68821-6>.

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

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2020