

Event-Related Potentials to Emotional Incongruence in Dogs Using Non-Invasive EEG

Hitomi Koike

Azabu University

Takatomi Kubo

`takatomi-k@is.naist.jp`

Nara Institute of Science and Technology

Yuki Maruno

Kansai University

Maaya Saito

Azabu University

Takumi Ozono

Azabu University

Natsuna Ishida

Nara Institute of Science and Technology

Kazunori Ohno

Tohoku University

Toshitaka Yamakawa

Quadlytics Inc.

Takefumi Kikusui

Azabu University

Miho Nagasawa

Azabu University

Article

Keywords:

Posted Date: April 2nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9133978/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on August 22nd, 2026.
See the published version at <https://doi.org/10.1038/s41598-026-66970-8>.

Event-Related Potentials to Emotional Incongruence in Dogs Using Non-Invasive EEG

Hitomi Koike¹, Takatomi Kubo^{2,*}, Yuki Maruno³, Maaya Saito¹, Takumi Ozono⁴, Natsuna Ishida², Kazunori Ohno⁴, Toshitaka Yamakawa^{2,5}, Takefumi Kikusui¹, and Miho Nagasawa¹

¹Faculty of Veterinary Medicine, Azabu University, Sagami-hara, 252-5201, Japan

²Graduate School of Science and Technology, Nara Institute of Science and Technology, Ikoma, 630-0192, Japan

³Faculty of Business Data Science, Kansai University, Suita, 565-8585, Japan

⁴Tough Cyber-physical AI Research Center, Tohoku University, Sendai, 980-0845, Japan

⁵Quadlytics Inc., Kyoto, 600-8813, Japan

*takatomi-k@is.naist.jp

ABSTRACT

Previous behavioral and autonomic studies suggest that dogs are sensitive to their human emotional expressions, yet the neural mechanisms underlying how such information is processed across modalities remain unclear. This study investigated whether dogs exhibit differential neural responses to emotional incongruence in their owner's multimodal expressions (face and voice) using electroencephalography (EEG). Dogs viewed audiovisual clips in which their owner's facial expression (happy or angry) was paired with a congruent or incongruent vocalization. Event-related potentials (ERPs) time-locked to voice onset were analyzed; incongruent pairings elicited significantly larger P300 and N400 responses at frontal sites than congruent pairings, a pattern consistent with mismatch detection and increased processing demands. These findings provide electrophysiological evidence that dogs integrate facial and vocal cues when processing their owners' emotional expressions, supporting multimodal emotion processing in the human–dog context. By revealing the temporal dynamics of this integration, the present study highlights multimodal processing of emotional information as a fundamental component of socio-emotional cognition, offering a comparative framework for investigations across species.

Introduction

Dogs are the earliest domesticated species, having coexisted with humans for approximately 30,000–40,000 years ago [1]. Over the course of this long relationship, behavioral experiments and autonomic nervous system indicators have suggested that dogs may recognize human emotions. In behavioral studies, dogs showed a tendency to look longer at combinations where a human facial expression and emotional voice were congruent, suggesting that dogs integrate information across different sensory modalities to discriminate between positive and negative emotions from humans [2]. Beyond behavior, physiological measures such as heart rate variability (HRV) have also been investigated. In a study by Katayama et al. (2019) [3], some dog–owner pairs showed positive correlations of R-R intervals and HRV indices under the emotional stress in the owners, indicating physiological synchrony between dogs and their owners.

At the neural level, an fMRI study demonstrated that in dogs, as in humans, a region near the auditory cortex responds more strongly to happy voice, indicating that dogs are neurally sensitive to vocal emotional valence expressions [4]. However, because fMRI provides limited temporal resolution for capturing the rapid dynamics of neural processing, our understanding of the neural mechanisms underpinning dogs' emotional processing remains limited. Just as the temporal dynamics of audiovisual emotional integration have been extensively investigated in humans using electroencephalography (EEG) (e.g., [5]), EEG offers a suitable approach for examining its millisecond-level dynamics in dogs.

Dogs' EEG research has progressed steadily since [6] established a non-invasive EEG recording protocol in awake, unrestrained dogs using positive reinforcement. Since then, event-related potentials (ERPs) have been reported in response to visual stimuli [6, 7] as well as auditory stimuli [8, 9], demonstrating that reliable neural responses can be obtained from awake dogs under non-invasive EEG recording conditions for both visual and auditory sensory modalities.

With respect to multimodal incongruence, dogs have been shown to understand object words given by humans, and an N400-like component is elicited when spoken label and visual object are mismatched [10], yet no studies have investigated whether dogs exhibit differential neural responses to emotional incongruence in multimodal expressions. Therefore, in the present study, we investigated neural responses in dogs while presenting emotionally congruent and incongruent pairings of their owner's facial and vocal expressions. By using ERPs, we aimed to capture the dynamics of emotional integration.

Human ERP studies have shown that emotional audiovisual incongruence modulates several ERP components. Mori et al. (2021) [11] demonstrated that the interaction between facial and vocal emotional expressions appears as a modulation of the auditory N400 amplitude in response to incongruent face–voice pairings. Moreover, enhanced N400 amplitudes for emotional incongruence have been observed across different emotional categories (happiness, neutrality, and anger), suggesting an automatic interaction of emotion related information across modalities [12].

Among the few ERP studies in dogs that have examined incongruence, Boros et al. (2024) [10] reported an N400-like effect when a spoken word was incongruent with a visual object. This suggests that dogs are also neurally sensitive to semantic incongruence. Based on these prior findings, we focused our primary analysis on the N400, where incongruence effects are typically observed.

Human studies have also examined earlier stages of audiovisual emotional integration. In particular, they have demonstrated that such integration occurs at early perceptual stages—within approximately 180 ms after voice onset—and that emotional incongruence modulates early ERP components [12, 13]. Ho et al. (2015) [14] reported that emotional incongruence was associated with reduced N100 amplitudes, whereas the P200 component exhibited bidirectional modulation depending on vocal emotion. However, findings on early components are not entirely consistent across studies, suggesting that depending on task demands, participants may flexibly switch between sensory-level and semantic-level integration [11, 15].

Dogs also exhibit early neural discrimination of emotional expressions: Kujala et al. (2020) [7] reported emotional differentiation occurring at approximately 130–170 ms, indicating that early perceptual sensitivity to emotional cues is not unique to humans. Based on this evidence, as potential indices of early multisensory integration, we examined the N100 (80–120 ms) and P200 (150–250 ms) components as supplementary analyses.

We hypothesized that emotional incongruence between facial and vocal expressions would elicit enhanced ERP responses, specifically increased N400 amplitudes reflecting the detection of emotional incongruence. By examining these ERP responses in dogs, this study aimed to investigate whether dogs exhibit differential neural responses to emotional incongruence in their owner’s multimodal expressions with high temporal resolution, thereby advancing our understanding of the neural basis of emotional cognition in dogs.

Results

To examine neural responses to emotional incongruence between owners’ facial and vocal expressions, we recorded non-invasive EEG while dogs viewed audiovisual stimuli. Electrodes were fixed using a custom-made EEG cap (Fig. 1a), and the auditory stimulus was synchronized with EEG via an auxiliary recording channel (Fig. 1b). The raw signals were filtered, cleaned using independent component analysis (ICA), and reconstructed for analysis (Fig. 1c). Seven electrodes (F3, F4, T3, T4, P3, P4, Cz) were used (Fig. 2). Dogs viewed randomized sequences of congruent and incongruent face–voice pairings (Fig. 3). ERPs were time-locked to voice onset and analyzed within predefined time windows corresponding to canonical N100, P200, P300, and N400 components.

Fig. 4 shows the grand-averaged ERP waveforms for the congruent and incongruent conditions across the three dogs. In the present study, we focus on the N400; however, in both conditions, the grand-averaged waveforms exhibited a positive-going deflection within this time window. Visually, this positivity could resemble a P300-like component, which reflects the allocation of attentional resources. In dogs, the polarity of the N400 is not consistently negative [16], making it difficult to dissociate a potential N400 from the sustained positivity associated with the P300. Therefore, we analyzed both the P300 (250–400 ms) and N400 (350–500 ms) time windows, interpreting any effects within these windows as reflective of emotional incongruence-related processing. In the approximately 250–500 ms range—corresponding to the predefined analysis windows for the P300 and N400 components—the amplitude of this positivity tended to be larger in the incongruent condition than in the congruent condition.

To statistically verify this trend, Welch’s two-sample t-tests were performed to compare the congruent and incongruent emotional conditions. For each ERP component (N100, P200, P300, N400), we examined whether significant differences were present and assessed the direction of amplitude changes.

1. t-test analysis

After averaging the ERP waveforms across the three dogs, comparisons revealed ERP components that are involved in the processing of emotional incongruence.

1.1 Group-level analysis

Group-level results are summarized in Table 2.

- N100 (80–120 ms) : Voltage values were significantly lower in the incongruent condition than in the congruent condition at Cz ($t = 2.33$, $p = .0198$).

- P200 (150-250 ms) : Voltage values were significantly higher in the incongruent condition at F3 ($t = -2.43$, $p = .0020$) and F4 ($t = -2.18$, $p = .0029$).
- P300 (250-400 ms) : Voltage values were significantly higher in the incongruent condition at F3 ($t = -3.84$, $p = .0001$).
- N400 (350-500 ms) : Voltage values were significantly higher in the incongruent condition at F3 ($t = -3.70$, $p = .0002$) and F4 ($t = -2.10$, $p = .0360$).

These results suggest that emotional incongruence influences both early (N100/P200) and later (P300/N400) stages of processing in dogs.

1.2 Individual-level analysis

Individual results are summarized in Table 3.

- Dog 1: significant N400 effect at F3 ($p < .05$)
- Dog 2: significant N400 effect at Cz ($p < .05$)
- Dog 3: significant effects for N100, P200, P300, and N400, each at ≥ 1 electrode ($p < .05$)

Notably, all three dogs showed a significant N400 incongruence effect, indicating consistent neural sensitivity to emotional mismatches across individuals. At the same time, these findings also revealed individual variability among dogs; therefore, a linear mixed-effects model (LMM) analysis was conducted to confirm whether the observed effects persisted after statistically accounting for inter-individual differences.

2. Linear mixed-effects model analysis

The LMM incorporated Dog as a random factor to account for inter-individual variability. The analysis confirmed that the interaction between Face and Voice remained significant for all electrode–component combinations that had shown significant effects in the Welch’s t-tests: P200, P300, and N400 at F3; P200 and N400 at F4; and N100 at Cz (all $p < .05$). This demonstrates that the differences between congruent and incongruent conditions persisted even after statistically controlling for individual variability among dogs.

Additionally, the main effect of Voice was significant at F3 across ERP components from N100 onward ($p < .05$), which aligns with the auditory stimulus-driven nature of the paradigm. However, because significant Face $\times$ Voice interactions were present, these main effects should be interpreted with caution as averaged effects across levels of the other factor. All statistical results, including fixed-effect coefficients and model fit indices, are provided in Table 4. Taken together, these LMM results converge with the t-test findings and confirm the modulation of ERP responses under emotional incongruence.

Discussion

The present study provides electrophysiological evidence that dogs exhibit differential neural responses to congruent and incongruent audiovisual emotional expressions from their owners, suggesting multimodal integration at the neural level. Importantly, this interpretation is supported by a clear P300-like response to the vocal stimuli, indicating attentional engagement with the owner’s voice; accordingly, the modulation of this response likely reflects facial–vocal congruency. Taken together, these findings support the interpretation that dogs integrate facial and vocal emotional information when processing their owners’ expressions.

Statistical analyses further showed that the observed modulation of ERP response differences across multiple ERP components driven by emotional congruence was statistically significant, and that the effect remained robust even after accounting for inter-individual variability in the mixed-effects model. At the individual level, significant N400 effects were confirmed in all dogs, suggesting that the modulation of ERP responses to audiovisual emotional incongruence was primarily driven by N400-related processes. Below, we discuss the potential functional roles of each ERP component (N100, P200, P300, and N400) in the processing of emotional incongruence.

With respect to semantic-level processing, the **N400** component is an ERP response linked to meaning processing and is generally observed as a larger negative deflection in response to semantically incongruent stimuli in humans [12, 17]. However, human ERP studies have shown that a frontal positivity often follows the typical N400 negativity [18], leading to an apparent polarity reversal at frontal electrode sites. Consistent with this, the significant difference between congruent and incongruent conditions observed at the frontal electrodes in the present study may represent a frontal N400-like effect associated with the detection of emotional incongruence.

In addition, emotional prime–target pairs have been shown to modulate the N400 depending on the duration of prosodic primes. Paulmann and Pell (2010) [19] reported that while a medium-length (400 ms) prosodic prime preceding a facial target elicited a typical priming effect—greater negativity for unrelated than for related pairs—a short (200 ms) prime reversed this pattern, producing greater negativity for related pairs. Although their paradigm presented the face after the voice, our vocal stimuli (1.65 s) corresponded to the short prosodic condition, suggesting that prosodic duration might have influenced the direction of the N400 modulation observed in the present study. In dog ERP studies, larger negative deflections have been reported between 220–470 ms for auditory sequences with low transitional probability [16], suggesting the presence of N400-like responses to expectancy violations in dogs, consistent with findings in humans. However, polarity differences have also been reported: when dogs were presented with spoken words paired with visual objects, a positive-going N400 mismatch effect was observed for incongruent object–word pairs [10]. Because scalp-recorded EEG reflects dipolar field activity, the polarity itself is not inherently meaningful; therefore, the positive-going deflection observed in our data can reasonably be interpreted as an N400-like response associated with the detection of semantic or emotional incongruence.

At a stage associated with attentional and contextual evaluation, the **P300** component reflects the allocation of attentional resources, while its amplitude is modulated by emotional and multimodal processing demands [20, 21, 22]. Larger P300 amplitudes have been interpreted as indexing increased cognitive workload, as the P300 is elicited when incoming stimuli require updates to an internal environmental model [23, 24, 25]. Further evidence shows that P300 amplitude and latency increase for incongruent conditions compared to congruent ones when visual letters and auditory sounds mismatch [26]. In the present experiment, a clear P300-like positivity was elicited by the audiovisual emotional stimuli, and this positivity was significantly larger for incongruent than congruent conditions. This pattern is consistent with the above interpretation that emotional conflict increases attentional and integrative processing demands. Although the primary incongruence effect in our study was reflected in differences in N400-related processes, the robust P300-like activity—as well as the possibility of latency prolongation—indicates that P300 modulation may also contribute to the differentiation between congruent and incongruent conditions.

At early stages of multisensory processing, the **N100** and **P200** components have been widely investigated in the context of audiovisual integration. Their amplitudes are typically suppressed during multisensory processing relative to unimodal stimulation, a pattern commonly interpreted as an index of facilitated audiovisual processing [27, 28, 29]. Both components have also been linked to emotional processing [30, 31]. However, the two components differ in functional specificity: whereas the P200 is sensitive to stimulus content and meaning and has been associated with the extraction of emotional significance [32], the N100 appears unaffected by audiovisual semantic incongruence and is instead modulated by predictability [14, 33, 34]. Early neural responses to visual and auditory stimuli in dogs have similarly been reported around 100 ms as well [6, 7, 8].

Regarding the present findings, the P200 exhibited a trend consistent with later components, a pattern that aligns with previous human ERP research reporting larger positive deflections for incongruent compared to congruent emotional trials at around 180 ms after stimulus onset [12]. In contrast, no significant incongruence effects were found for the N100 at frontal sites. This pattern suggests that semantically driven audiovisual emotional integration in dogs may emerge after the N100 time window. Nonetheless, at Cz, a significant increase in N100 amplitude was observed, suggesting that an early stage of emotional multisensory integration cannot be fully ruled out.

Furthermore, the enhanced P200 amplitude under incongruent conditions may be interpreted within predictive-processing accounts as reflecting additional processing demands during multimodal (audiovisual) integration when cross-modal expectations are not met. However, the opposite pattern—P200 suppression for incongruent emotional stimuli—has also been documented in humans [35], and Knowland et al. (2014) [29] further suggested that such audiovisual P200 suppression may reflect competition between mismatching multisensory inputs. These divergent findings underscore the need for caution when interpreting P200 modulations.

Taken together, although N100 and P200 modulations under audiovisual emotional incongruence have been investigated in humans, the reported patterns remain inconsistent across studies [11, 14, 15, 36, 37]. One proposed explanation is that humans may flexibly engage either semantic-level or sensory-level processing depending on task demands or stimulus features [11]. If such flexible switching also exists in dogs—and under what conditions—remains an open question requiring further investigation.

In addition to congruence-related effects, we also explored whether emotional valence might modulate ERP responses, although any interpretation should be considered preliminary. LMM analysis revealed a significant interaction at electrode F3, whereas the main effects for happy face and happy voice showed positive coefficients ($\beta > 0$), indicating a numerical tendency toward larger ERP amplitudes in the happy condition. These estimates were obtained in the presence of a significant interaction and should therefore be interpreted with caution, as the apparent increase is likely to reflect incongruence-related effects rather than a main effect of valence. Moreover, this tentative pattern differs from some human ERP findings reporting larger amplitudes for negative relative to positive emotions [38, 39], further underscoring the need for caution in interpreting valence-related effects in the present study.

However, several limitations should be acknowledged. First, although the number of recording sessions did not differ substantially across dogs, the number of trials retained after artefact rejection varied between individuals. Dog3 contributed a larger number of analyzable trials. This imbalance may have increased the statistical weight of Dog3 in the individual-level analyses and the linear mixed-effects models. Consequently, the observed effect sizes may have been influenced by unequal trial numbers. Although random effects were included to account for individual variability, unequal trial numbers remain a methodological limitation. In addition, it remains challenging to clearly disentangle the specific contributions of individual ERP components. Higher-resolution approaches, such as source localization, time–frequency analyses, or multivariate decoding, may help clarify the neural mechanisms underlying emotional integration.

Nevertheless, the present study also has important strengths. The main findings were supported by converging results from both Welch's t-tests and linear mixed-effects models, indicating that the observed effects were robust across different analytical approaches. Moreover, the study extends canine ERP research beyond exclusively positive emotional conditions. In canine EEG experiments, maintaining stillness during recording is critical, as even minimal muscle activity can introduce electromyographic noise. Consequently, negatively valenced stimuli have often been avoided to minimize arousal-related movement artifacts [9]. By successfully incorporating both positive and negative emotional expressions, the present study represents a methodological advancement in canine ERP research and broadens the scope of emotional processing paradigms that can be investigated using EEG in dogs.

Future research should examine a broader range of emotional combinations to further elucidate the neural mechanisms underlying dogs' perception of human emotions. In addition, incorporating emotional stimuli from unfamiliar individuals may help determine whether such neural responses depend on familiarity and experience with the owner, or whether they reflect a more generalized, evolutionarily conserved mechanism. Beyond the audiovisual domain, dogs are also known to recognize human emotional states through olfactory cues [40]. Thus, future studies combining EEG with olfactory stimuli may allow high–temporal-resolution investigations of emotion processing in the olfactory modality and further clarify the associated neural mechanisms. Against this background, EEG-based approaches offer a powerful tool for investigating dogs' emotional perception with high temporal precision. Importantly, even when animals exhibit similar behaviors, the underlying neural processes may not be identical [41]. Accordingly, combining behavioral and electrophysiological approaches will be essential for a comprehensive understanding of the mechanisms supporting emotional perception.

In conclusion, the present study provides electrophysiological evidence that dogs process their owners' emotional expressions in a multimodal manner. To our knowledge, this is the first study to suggest that incongruence between facial and vocal emotional cues elicited systematic neural modulations, particularly in late ERP components, compared with congruent pairings. These neural differences are consistent with mismatch detection and increased processing demands, suggesting that dogs do not merely respond to isolated sensory cues but integrate information across modalities when interpreting their owners' emotions. Importantly, the observed effects were consistent across individuals, despite variability in signal magnitude, indicating a shared neural sensitivity to multimodal emotional incongruence. Although the small sample size warrants caution in generalization, the convergence of effects across dogs supports the interpretation that multimodal processing of emotional expressions constitutes a fundamental aspect of human–dog socio-emotional communication. By providing millisecond-level access to the temporal dynamics of such processing, ERP-based approaches offer a valuable framework for future comparative studies of emotional cognition across species.

Methods

Ethics statement.

All experimental procedures involving animals were approved by the Animal Ethics Committee of Azabu University (#180410-1) and were performed in accordance with relevant institutional and national guidelines for animal care and use. The study complied with the ARRIVE guidelines for reporting animal research. Procedures involving human participants were approved by the Ethical Committee for Medical and Health Research Involving Human Subjects of Azabu University (#052) and were conducted in accordance with relevant guidelines and regulations, including the Declaration of Helsinki. The dogs were privately owned pets. Written informed consent was obtained from all dog owners for both their dogs' participation and their own participation in the study. In addition, informed consent was obtained from all human participants who provided emotional ratings of the stimuli, and they were informed that they could withdraw their consent at any time without penalty. Consent was also obtained from the participant whose images are shown in the figures for publication of identifiable images in this open-access journal.

Subjects and Experimental setup

This study was performed in a laboratory at the Department of Veterinary Medicine at Azabu University. The experimental room (1.6 m × 4.6 m) minimized electrical interference and provided optimal conditions for the experiment. EEG recordings

were performed on three healthy dogs (Dog1: a 6-year-old male Brittany Spaniel, Dog2: a 10-year-old male Standard Poodle, Dog3: a 5-year-old female Standard Poodle) that were trained to remain completely still while lying on a measurement mattress (Fig. 1a).

The dogs were also conditioned to tolerate wearing an EEG cap designed to fix the electrodes (AP-P130(A)-08, Miyuki Giken Co., Ltd., Tokyo, Japan) in place, rest their heads on a chin rest, and remain stationary in front of the stimulus-presentation monitor (ASUS ZenScreen MB229CF-J, Japan) screen during recording. Although dogs were not physically restrained during the task, positive reinforcement was provided to encourage them to remain stationary.

During the sessions, the dogs rested on the experimental platform with cushions elevated to a height of 0.6 m. A computer (Dell Inc., Intel Core i3-1115G4 3.00GHz, 8GB RAM, Windows 11 Pro 64-bit, ver. 24H2; Dell Inc., Round Rock, TX, USA) for EEG signal recording was positioned adjacent to the platform, and the EEG amplifier was positioned either directly or in close proximity to the body. A video camera (JVC Everio R GZ-F200-T, JVC KENWOOD Corporation, Yokohama, Japan) was mounted on top of the monitor to record the dogs' behavior during the experiment. A speaker (SRS-XG300, Sony Corporation, Tokyo, Japan) that emitted auditory stimuli was placed behind the computer to control the stimuli. Dog owners remained outside the experimental room throughout the sessions, while two investigators were positioned beside and behind each dog. All dogs were privately owned pets, and informed consent for their participation was obtained from their owners.

EEG measurement

The dogs were given time (approximately 5 min) to acclimate to the experimental room before the recording began to ensure voluntary participation and minimize stress. Once relaxed, the dogs were guided to a platform facing the stimulus presentation monitor, and the electrodes were carefully affixed to their heads. After confirming the setup, the investigators initiated the data acquisition and presentation of the stimuli, standing unobtrusively in a location outside the dog's field of view.

Several preparatory steps were performed to optimize the contact between the electrodes and the skin. First, the dogs' heads were wiped with alcohol swabs to remove skin oils from the electrode sites. Then, Neurgel (NEURGEL250F-6, Atris Pty Ltd., Sydney, Australia) was applied directly to the electrode sites to enhance the electrical conductivity and ensure stable attachment. Finally, Custom-made EEG caps designed to accommodate the head anatomy of each dog were fitted to maintain the position of the electrodes during the recordings. EEG signals were recorded using a Polymate Pocket (MP208; Miyuki Giken Co., Ltd., Tokyo, Japan), and the data were preprocessed using Bio Signal Viewer software (Miyuki Giken Co., Ltd.). The signal was digitized at a sample rate of 1000 Hz, and the built-in hardware band-pass filter was set to 0.53-30 Hz during acquisition. To ensure precise synchronization between the audiovisual stimuli and the EEG recordings, the audio signal used as the auditory stimulus was simultaneously fed into an auxiliary EEG channel. This allowed us to define the ERP time-locking point directly from the recorded audio waveform rather than relying solely on software triggers (Fig. 1b).

The electrode configuration included seven recording sites (F3, F4, T3, T4, P3, P4, Cz) strategically positioned to capture neural activity across key cortical regions (Fig. 2). Electrodes F3 and F4 were placed bilaterally over the frontal cortex, and electrodes T3 and T4 were positioned laterally over the temporal cortex. Electrodes P3 and P4 were situated in the parieto-occipital region, and electrode Cz was located along the midline. The occipital protuberance at the rear of the skull served as the site for the reference electrode and the ground electrode was affixed near the base of the left ear. This electrode configuration was derived from the methodology outlined by Kujala et al. (2020) [7]. Electrode impedance was confirmed to be below 10 k Ω and at least below 20 k Ω prior to recording to ensure high-fidelity data acquisition. To maintain high-quality EEG recordings, both electrode stability and the dog's gaze and attentional state were carefully and continuously monitored throughout the experiment. Only trials in which the dog was visually attending to the screen were included in the analysis. The dog's overall condition was also monitored to detect any signs of fatigue or stress. If the dog showed any indication of fatigue, the session was immediately terminated to ensure its well-being. Depending on the dog's training level and physical condition, five to ten sessions were conducted per day.

Experimental Procedure

Dogs were presented with audiovisual clips depicting their owner's emotions (Fig. 3). ERP amplitude has been shown to be associated with emotional arousal [42]; to ensure comparable arousal levels across emotional conditions, we selected happiness and anger as the owner's emotional expressions. The visual stimuli were static photographs of the owner's face, depicting either emotion. Auditory stimuli were amplitude-normalized recordings of the same word (e.g., "really?"), commonly used by the owner in daily life, and were duration-matched across emotional conditions. This design controlled for lexical content while systematically varying emotional valence. Prior to the experiment, 28 human participants were asked to rate the emotional valence of all facial and vocal stimuli used in the study. All participants provided informed consent prior to participation. Emotional valence ratings were then analyzed using the Mann-Whitney U test, confirming significant differences between positive and negative stimuli across all items ($p < .001$).

Each dog viewed a single 68-s video sequence comprising 12 trials, preceded by an 8-s background-only period to stabilize EEG activity. In each trial, a neutral background image was continuously displayed; the owner's facial photograph appeared for 3 s, and the auditory stimulus was presented 1.5 s after face onset (SOA = 1.5 s). Congruent (face-voice match) and incongruent (face-voice mismatch) pairings were counterbalanced across trials. To avoid potential bias in the presentation order, the sequence of stimulus combinations was randomized using Python when generating the video. Between trials, a 2-s background image was shown to allow the EEG signal to return toward baseline. The 1.5-s SOA was aimed to avoid overlap between early visually evoked components (P100, N100, P200) following face onset and the auditory evoked response, thereby isolating the ERP elicited by the voice. Visual activity can persist for several hundred milliseconds after onset; delaying the voice until at least 1 s reduces temporal overlap. At the same time, the SOA was kept within the upper limit of the audiovisual integration window (≈ 2 s), beyond which the events are more likely to be processed as separate rather than integrated [43].

EEG data processing and artefact rejection

Offline preprocessing was performed using Python (version 3.11). Continuous EEG data were segmented into epochs ranging from -1500 ms to +1500 ms relative to the detected audio voice onset on the auxiliary channel. The data were notch-filtered at 50 Hz to remove line noise and then band-pass filtered between 0.05 and 40 Hz using a zero-phase second-order sections (sos) filter. Any epoch in which the EEG signal exceeded $\pm 100 \mu\text{V}$ at any electrode was automatically rejected as containing artifacts. After filtering, Independent Component Analysis (ICA) was applied to remove non-neural artifacts. EEG data from each session were decomposed into eight independent components using the FastICA algorithm (scikit-learn, Python 3.11). Each component was visually inspected while simultaneously referencing the synchronized video recordings of the dog. Components clearly associated with artifacts-including eye movements, muscle activity, and the audio synchronization channel-were removed. The remaining components were then back-projected to reconstruct artifact-free EEG signals.

In this way, artifacts were removed and only reconstructed clean EEG signals were used for subsequent analyses (Fig. 1c). After artifact rejection, the mean number of remaining trials per dog ranged from 100 to 434 across the four conditions; on average, 487 trials were retained for congruent conditions and 500 for incongruent conditions. Although the number of retained trials was balanced between congruent and incongruent conditions, inter-individual variability in data quantity required statistical adjustment in the subsequent modeling. The number of trials analyzed for each individual and condition is summarized in Table 1.

Statistical analysis

Statistical analyses were performed using MATLAB (R2023b). For each ERP component, Welch's two-sample t-tests and linear mixed-effects models (LMMs) were applied within time windows commonly reported in previous studies to examine the effects of congruency. Welch's t-tests were used because they do not assume equal variances between conditions and are robust to potential differences in variance across trial distributions. These complementary approaches were combined to assess whether the observed effects remained consistent after accounting for individual variability. Analyses were conducted on epochs time-locked to the onset of the auditory stimulus. In the LMM, individual dogs were included as a random intercept to account for between-subject variability. Given the limited number of subjects ($n=3$), random slope estimation was considered unreliable and therefore was not included in the model. To further characterize within-subject patterns, complementary Welch's two-sample t-tests were conducted separately for each individual, and individual trends were qualitatively described.

Previous neuroimaging research has shown that interactions between visual and auditory modalities for emotional information do not occur in primary sensory cortices; instead, emotional cues from faces and prosody are integrated later in higher-order associative cortices and limbic regions [44]. Furthermore, emotional incongruence between visual and auditory signals has been shown to activate the cingulate-frontoparietal network implicated in conflict monitoring and resolution [45]. Consistent with these findings, ERP studies in humans have reported significant facial-vocal congruency effects over frontal and frontocentral regions [14]. Based on this converging evidence, the present study focused on electrodes corresponding to the frontal-parietal regions (F3, F4, Cz).

Four ERP components associated with audiovisual emotional incongruence-N100 (80-120 ms), P200 (150-250 ms), P300 (250-400 ms), and N400 (350-500 ms)-were extracted, and amplitude differences between congruent and incongruent conditions were analyzed for each component. Time windows for each ERP component were defined based on ranges commonly reported in human studies, while taking into account that these windows vary across prior research and adjusting them to suit the present experimental conditions. Reconstructed EEG data, which had been segmented time-locked to the auditory stimulus, were labeled for each dog at electrodes F3, F4, and Cz according to the four face-voice combinations. All sessions remaining after artifact rejection were included in the analysis. Within the defined time windows for each ERP component, data were averaged for each combination of individual, condition, and electrode.

1. Primary Analysis: t-tests

Initially, we examined whether there were significant differences in the averaged ERP waveforms between congruent and incongruent conditions at three electrodes (F3, F4, Cz) across all three dogs. Welch's two-sample t-tests were conducted for each ERP component (N100, P200, P300, and N400). Given the potential for inter-individual variability in dogs, both group-level and individual-level t-tests were performed to evaluate overall and subject-specific tendencies, respectively.

2. Supplementary Analysis: Linear Mixed-effects Model

To further account for inter-individual variability and to decompose the effects into main and interaction components, a linear mixed-effects model (LMM) was applied to the same comparisons. Because ERP amplitudes were analyzed at the trial level, the LMM additionally accounted for the non-independence of trials nested within subjects. The model included facial emotion and vocal emotion as fixed effects, along with their interaction term, and incorporated subject identity as a random effect. However, because the number of subjects was limited to three, it should be noted that the stability of random-effect estimation may be limited. Therefore, this analysis served to examine whether the primary t-test findings remained robust after accounting for individual variability.

References

1. Vilà, C. *et al.* Multiple and ancient origins of the domestic dog. *Science* **276**, 1687–1689, DOI: <https://doi.org/10.1126/science.276.5319.1687> (1997).
2. Albuquerque, N. *et al.* Dogs recognize dog and human emotions. *Biol. Lett.* **12**, 20150883, DOI: <https://doi.org/10.1098/rsbl.2015.0883> (2016).
3. Katayama, M. *et al.* Emotional contagion from humans to dogs is facilitated by duration of ownership. *Front. Psychol.* **10**, 1678, DOI: <https://doi.org/10.3389/fpsyg.2019.01678> (2019).
4. Andics, A., Gácsi, M., Faragó, T., Kis, A. & Miklósi, A. Voice-sensitive regions in the dog and human brain are revealed by comparative fMRI. *Curr. Biol.* **24**, 574–578, DOI: <https://doi.org/10.1016/j.cub.2014.01.058> (2014).
5. Gao, C. *et al.* Temporal dynamics of audiovisual affective processing. *Biol. Psychol.* **139**, 59–72, DOI: <https://doi.org/10.1016/j.biopsycho.2018.10.001> (2018).
6. Törnqvist, H. *et al.* Visual event-related potentials of dogs: a non-invasive electroencephalography study. *Anim. Cogn.* **16**, 973–982, DOI: <https://doi.org/10.1007/s10071-013-0630-2> (2013).
7. Kujala, M. V. *et al.* Time-resolved classification of dog brain signals reveals early processing of faces, species and emotion. *Sci. Rep.* **10**, 19846, DOI: <https://doi.org/10.1038/s41598-020-76806-8> (2020).
8. Magyari, L., Huszár, Z., Turzó, A. & Andics, A. Event-related potentials reveal limited readiness to access phonetic details during word processing in dogs. *R. Soc. Open Sci.* **7**, 200851, DOI: <https://doi.org/10.1098/rsos.200851> (2020).
9. Bálint, A., Eleőd, H., Magyari, L., Kis, A. & Gácsi, M. Differences in dogs' event-related potentials in response to human and dog vocal stimuli; a non-invasive study. *R. Soc. Open Sci.* **9**, 211769, DOI: <https://doi.org/10.1098/rsos.211769> (2022).
10. Boros, M. *et al.* Neural evidence for referential understanding of object words in dogs. *Curr. Biol.* **34**, 1750–1754.e4, DOI: <https://doi.org/10.1016/j.cub.2024.02.029> (2024).
11. Mori, K., Tanaka, A., Kawabata, H. & Arao, H. The N400 and late occipital positivity in processing dynamic facial expressions with natural emotional voice. *Neuroreport* **32**, 858–863, DOI: <https://doi.org/10.1097/WNR.0000000000001669> (2021).
12. Föcker, J. & Röder, B. Event-related potentials reveal evidence for late integration of emotional prosody and facial expression in dynamic stimuli: An ERP study. *Multisens. Res.* **32**, 473–497, DOI: <https://doi.org/10.1163/22134808-20191332> (2019).
13. de Gelder, B., Böcker, K. B., Tuomainen, J., Hensen, M. & Vroomen, J. The combined perception of emotion from voice and face: early interaction revealed by human electric brain responses. *Neurosci. Lett.* **260**, 133–136, DOI: [https://doi.org/10.1016/s0304-3940\(98\)00963-x](https://doi.org/10.1016/s0304-3940(98)00963-x) (1999).
14. Ho, H. T., Schröger, E. & Kotz, S. A. Selective attention modulates early human evoked potentials during emotional face-voice processing. *J. Cogn. Neurosci.* **27**, 798–818, DOI: https://doi.org/10.1162/jocn_a_00734 (2015).

15. Garrido-Vásquez, P., Pell, M. D., Paulmann, S. & Kotz, S. A. Dynamic facial expressions prime the processing of emotional prosody. *Front. Hum. Neurosci.* **12**, 244, DOI: <https://doi.org/10.3389/fnhum.2018.00244> (2018).
16. Boros, M. *et al.* Neural processes underlying statistical learning for speech segmentation in dogs. *Curr. Biol.* **31**, 5512–5521.e5, DOI: <https://doi.org/10.1016/j.cub.2021.10.017> (2021).
17. Kutas, M. & Federmeier, K. D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* **62**, 621–647, DOI: <https://doi.org/10.1146/annurev.psych.093008.131123> (2011).
18. Thornhill, D. E. & Van Petten, C. Lexical versus conceptual anticipation during sentence processing: frontal positivity and N400 ERP components. *Int. J. Psychophysiol.* **83**, 382–392, DOI: <https://doi.org/10.1016/j.ijpsycho.2011.12.007> (2012).
19. Paulmann, S. & Pell, M. D. Contextual influences of emotional speech prosody on face processing: how much is enough? *Cogn. Affect. Behav. Neurosci.* **10**, 230–242, DOI: <https://doi.org/10.3758/CABN.10.2.230> (2010).
20. Polich, J. Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* **118**, 2128–2148, DOI: <https://doi.org/10.1016/j.clinph.2007.04.019> (2007).
21. Liu, T. *et al.* Emotional cues during simultaneous face and voice processing: electrophysiological insights. *PLoS One* **7**, e31001, DOI: <https://doi.org/10.1371/journal.pone.0031001> (2012).
22. Lu, Z., Li, Q., Gao, N., Yang, J. & Bai, O. Happy emotion cognition of bimodal audiovisual stimuli optimizes the performance of the P300 speller. *Brain Behav.* **9**, e01479, DOI: <https://doi.org/10.1002/brb3.1479> (2019).
23. Kok, A. On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology* **38**, 557–577, DOI: <https://doi.org/10.1017/s0048577201990559> (2001).
24. Donchin, E. & Coles, M. G. H. Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* **11**, 357–374, DOI: <https://doi.org/10.1017/s0140525x00058027> (1988).
25. Wickens, C., Kramer, A., Vanasse, L. & Donchin, E. Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science* **221**, 1080–1082, DOI: <https://doi.org/10.1126/science.6879207> (1983).
26. Andres, A. J. D., Oram Cardy, J. E. & Joanisse, M. F. Congruency of auditory sounds and visual letters modulates mismatch negativity and P300 event-related potentials. *Int. J. Psychophysiol.* **79**, 137–146, DOI: <https://doi.org/10.1016/j.ijpsycho.2010.09.012> (2011).
27. Klucharev, V., Möttönen, R. & Sams, M. Electrophysiological indicators of phonetic and non-phonetic multisensory interactions during audiovisual speech perception. *Brain Res. Cogn. Brain Res.* **18**, 65–75, DOI: <https://doi.org/10.1016/j.cogbrainres.2003.09.004> (2003).
28. van Wassenhove, V., Grant, K. W. & Poeppel, D. Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 1181–1186, DOI: <https://doi.org/10.1073/pnas.0408949102> (2005).
29. Knowland, V. C. P., Mercure, E., Karmiloff-Smith, A., Dick, F. & Thomas, M. S. C. Audio-visual speech perception: a developmental ERP investigation. *Dev. Sci.* **17**, 110–124, DOI: <https://doi.org/10.1111/desc.12098> (2014).
30. Scott, G. G., O'Donnell, P. J., Leuthold, H. & Sereno, S. C. Early emotion word processing: evidence from event-related potentials. *Biol. Psychol.* **80**, 95–104, DOI: <https://doi.org/10.1016/j.biopsycho.2008.03.010> (2009).
31. Paulmann, S., Jessen, S. & Kotz, S. A. Investigating the multimodal nature of human communication: Insights from ERPs. *J. Psychophysiol.* **23**, 63–76, DOI: <https://doi.org/10.1027/0269-8803.23.2.63> (2009).
32. Michalski, A. Expectation of an important event affects responses to irrelevant stimuli of different modalities. *Acta Neurobiol. Exp. (Wars.)* **60**, 467–478, DOI: <https://doi.org/10.55782/ane-2000-1366> (2000).
33. Vroomen, J. & Stekelenburg, J. J. Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J. Cogn. Neurosci.* **22**, 1583–1596, DOI: <https://doi.org/10.1162/jocn.2009.21308> (2010).
34. Stekelenburg, J. J. & Vroomen, J. Electrophysiological correlates of predictive coding of auditory location in the perception of natural audiovisual events. *Front. Integr. Neurosci.* **6**, 26, DOI: <https://doi.org/10.3389/fnint.2012.00026> (2012).

35. Kokinous, J., Kotz, S. A., Tavano, A. & Schröger, E. The role of emotion in dynamic audiovisual integration of faces and voices. *Soc. Cogn. Affect. Neurosci.* **10**, 713–720, DOI: <https://doi.org/10.1093/scan/nsu105> (2015).
36. Zinchenko, A., Obermeier, C., Kanske, P., Schröger, E. & Kotz, S. A. Positive emotion impedes emotional but not cognitive conflict processing. *Cogn. Affect. Behav. Neurosci.* **17**, 665–677, DOI: <https://doi.org/10.3758/s13415-017-0504-1> (2017).
37. Pourtois, G., de Gelder, B., Vroomen, J., Rossion, B. & Crommelinck, M. The time-course of intermodal binding between seeing and hearing affective information. *Neuroreport* **11**, 1329–1333, DOI: <https://doi.org/10.1097/00001756-200004270-00036> (2000).
38. Gardener, E. K. T., Carr, A. R., Macgregor, A. & Felmingham, K. L. Sex differences and emotion regulation: an event-related potential study. *PLoS One* **8**, e73475, DOI: <https://doi.org/10.1371/journal.pone.0073475> (2013).
39. Lithari, C. *et al.* Are females more responsive to emotional stimuli? a neurophysiological study across arousal and valence dimensions. *Brain Topogr.* **23**, 27–40, DOI: <https://doi.org/10.1007/s10548-009-0130-5> (2010).
40. D’Aniello, B., Semin, G. R., Alterisio, A., Aria, M. & Scandurra, A. Interspecies transmission of emotional information via chemosignals: from humans to dogs (*canis lupus familiaris*). *Anim. Cogn.* **21**, 67–78, DOI: <https://doi.org/10.1007/s10071-017-1139-x> (2018).
41. Fitch, W. T., Huber, L. & Bugnyar, T. Social cognition and the evolution of language: constructing cognitive phylogenies. *Neuron* **65**, 795–814, DOI: <https://doi.org/10.1016/j.neuron.2010.03.011> (2010).
42. Li, Y., Zhang, M., Liu, S. & Luo, W. EEG decoding of multidimensional information from emotional faces. *Neuroimage* **258**, 119374, DOI: <https://doi.org/10.1016/j.neuroimage.2022.119374> (2022).
43. van Wassenhove, V., Grant, K. W. & Poeppel, D. Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* **45**, 598–607, DOI: <https://doi.org/10.1016/j.neuropsychologia.2006.01.001> (2007).
44. Chen, Y.-H. *et al.* Neuromagnetic oscillations to emotional faces and prosody: Neuromagnetic oscillations to emotional faces and prosody. *Eur. J. Neurosci.* **31**, 1818–1827, DOI: <https://doi.org/10.1111/j.1460-9568.2010.07203.x> (2010).
45. Müller, V. I. *et al.* Incongruence effects in crossmodal emotional integration. *Neuroimage* **54**, 2257–2266, DOI: <https://doi.org/10.1016/j.neuroimage.2010.10.047> (2011).

Acknowledgements

We thank Animal Life Solutions Co., Ltd., especially Takuma Suzuki, for assisting with dog training and measurements.

Author contributions statement

H.K. designed the methodology, collected the data, performed the data analysis, and drafted the manuscript. T.Ku. developed the analytical methods and contributed to the methodology and manuscript revision. Y.M. wrote the code for the data analysis and for the randomised stimulus presentation videos. M.S. assisted with the experiments, dog training, and methodological procedures. T.O. constructed the EEG caps. N.I. contributed to the development of the data processing methods. K.O. acquired funding and provided methodological advice. T.Y. provided methodological advice. T.Kik. provided methodological advice. M.N. acquired funding and provided advice. All authors reviewed and approved the final manuscript.

Funding

This study was supported by the Japan Science and Technology Agency (JST), CREST (Grant Number JPMJCR22P2).

Additional information

Data availability. The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request. **Competing interests.** The authors declare no competing interests.

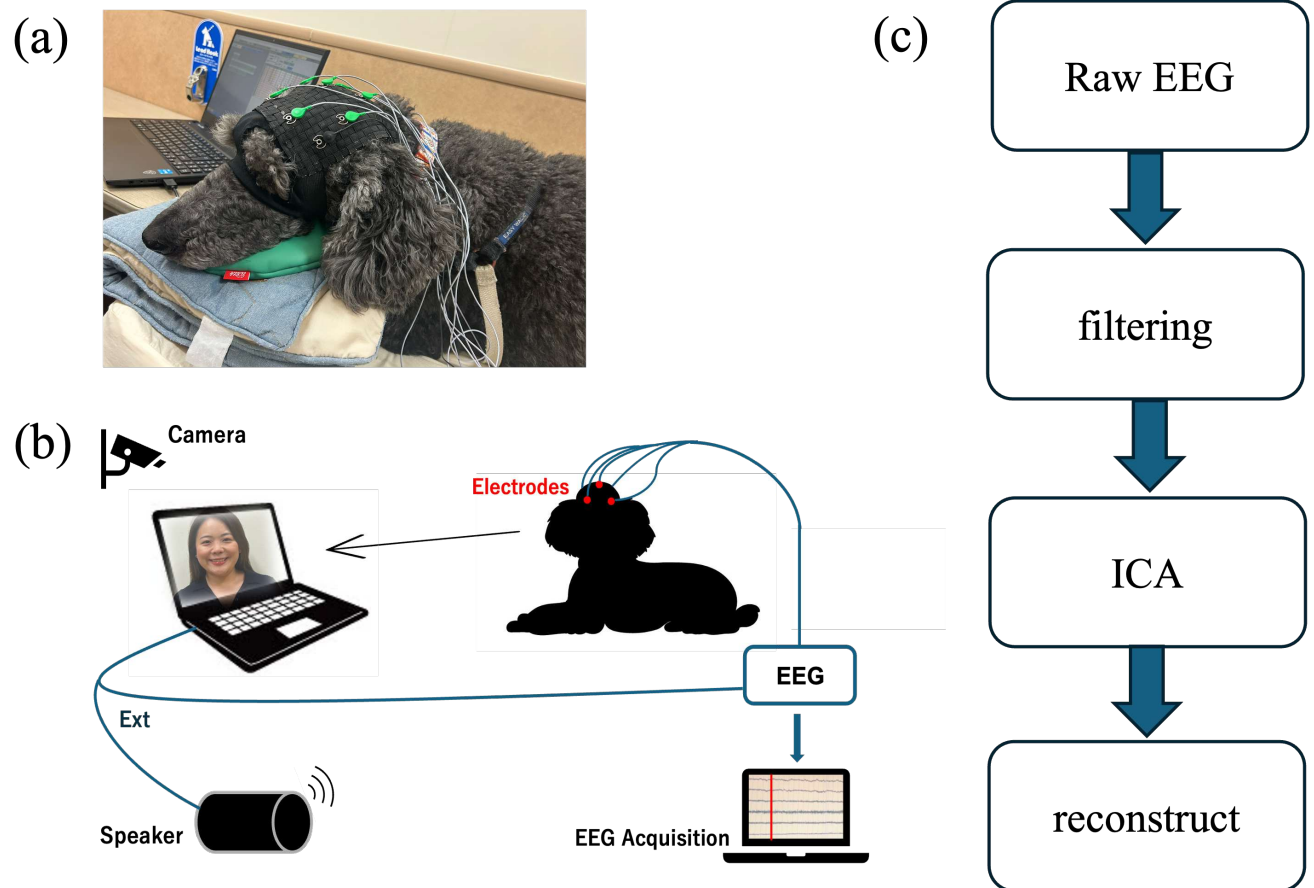


Figure 1. Experimental setup and EEG preprocessing pipeline. (a) Photograph of Dog 2 during EEG recording with the custom-made EEG cap. (b) Schematic overview of the experimental setup showing the audiovisual stimulus presentation and EEG acquisition system. The external trigger channel (Ext) was used to synchronize auditory stimulus onset with EEG recording. (c) EEG preprocessing pipeline including filtering, independent component analysis (ICA), and signal reconstruction after artefact removal.

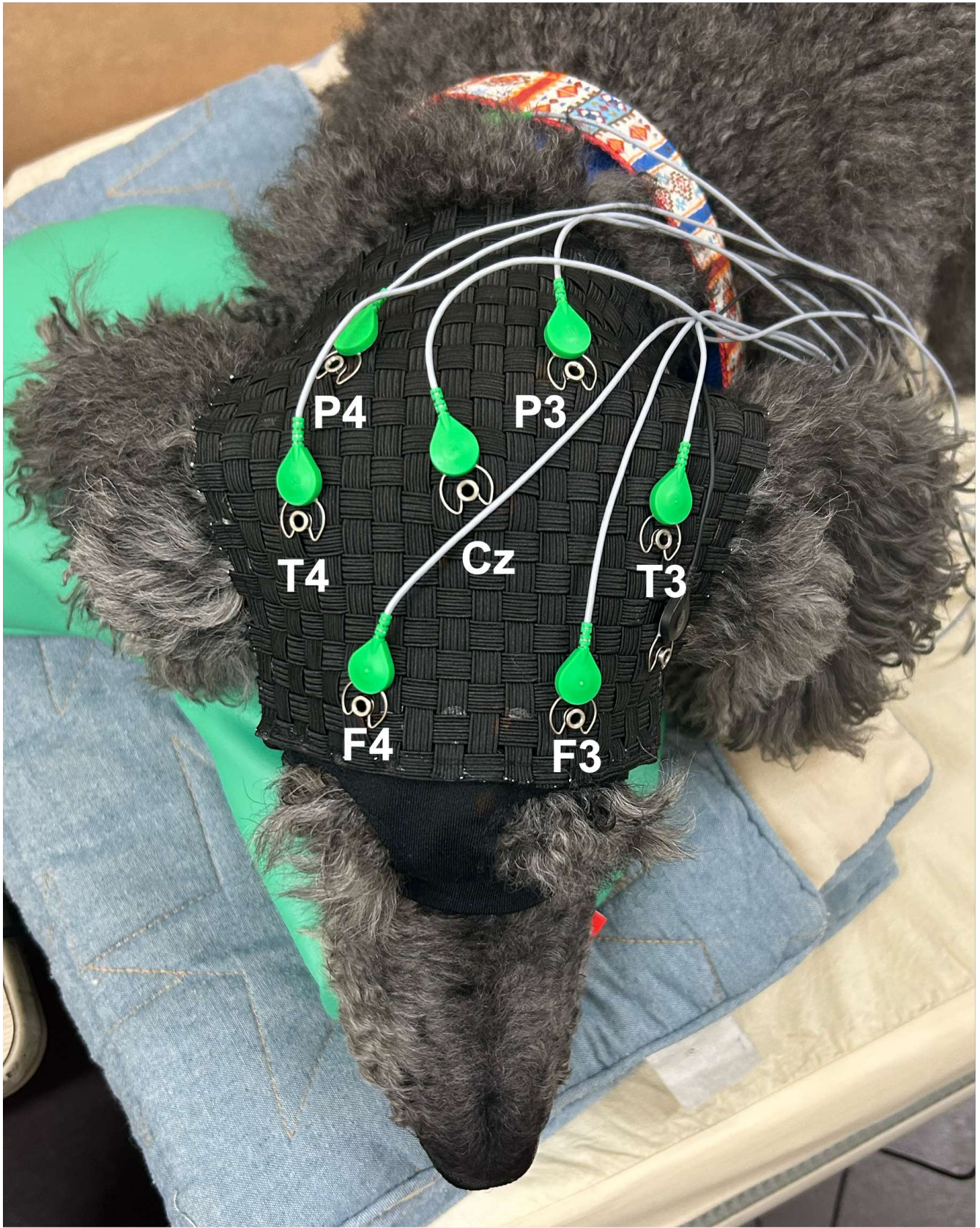


Figure 2. Electrode configuration used for EEG recording in dogs. Electrode positions (F3, F4, T3, T4, P3, P4, Cz) fixed with a custom-made EEG cap during recording.

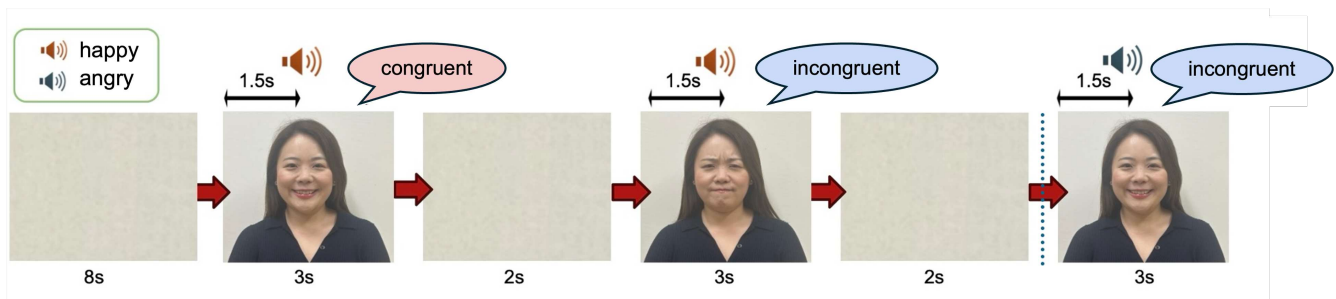


Figure 3. Timeline of audiovisual stimulus presentation. Each sequence began with an 8-s baseline period followed by trials consisting of a facial stimulus (3 s). The vocal stimulus was presented 1.5 s after face onset, creating congruent or incongruent combinations of facial and vocal emotional expressions. Each video sequence contained 12 trials with equal numbers of congruent and incongruent conditions. Written informed consent for the publication of these images was obtained from the participant.

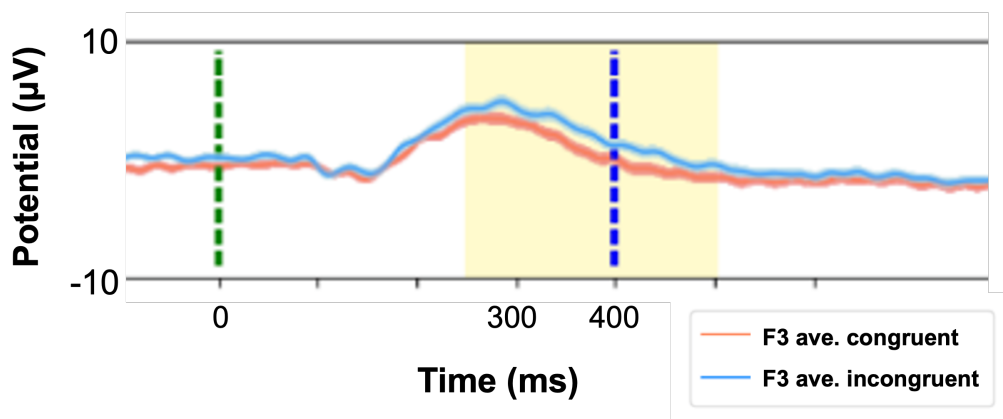


Figure 4. Grand-averaged ERP waveforms across the three dogs at the F3 electrode for congruent and incongruent conditions. Red line: congruent condition; blue line: incongruent condition. The vocal stimulus occurred at 0 ms (vertical green dashed line). The blue dashed line marks 400 ms after voice onset. The yellow shaded region highlights the analysis window corresponding to the P300 and N400 components (250-500 ms).

Dog ID	Congruency	Condition (face/voice)	Trial count	Total
1	Congruent	angry/angry happy/happy	99 102	201
	Incongruent	angry/happy happy/angry	98 101	199
2	Congruent	angry/angry happy/happy	197 211	408
	Incongruent	angry/happy happy/angry	211 207	418
3	Congruent	angry/angry happy/happy	412 440	852
	Incongruent	angry/happy happy/angry	432 452	884

Table 1. Number of remaining and analysed trials for each dog in the four conditions after artefact rejection.

ERPtype	Channel	Mean_Mismatch	<i>t</i>	df	<i>p</i>
N100	F3	-0.973	-1.43	2946.2	0.153
	F4	-1.36	0.0437	2959.7	0.965
	Cz	-1.28	2.33	2884.1	0.0198*
P200	F3	3.00	-2.43	2957.7	0.0152*
	F4	1.91	-2.18	2924.6	0.0291*
	Cz	0.969	0.287	2927.9	0.774
P300	F3	6.18	-3.84	2958.1	0.000127***
	F4	4.30	-1.62	2939.5	0.106
	Cz	1.53	-0.431	2901.7	0.666
N400	F3	1.25	-3.70	2928.7	0.000222***
	F4	0.482	-2.10	2941.0	0.0360*
	Cz	-0.0499	-0.972	2956.3	0.331

Table 2. Detailed results of group-level Welch's two-sample t-tests. Two-tailed tests were used. Significant effects are shown in bold, with significance levels indicated by asterisks (* $p < .05$, ** $p < .01$, *** $p < .001$).

ERPtype	Channel	Dog 1			Dog 2			Dog 3		
		<i>t</i>	df	<i>p</i>	<i>t</i>	df	<i>p</i>	<i>t</i>	df	<i>p</i>
N100	F3	-0.812	398.0	0.417	1.94	823.7	0.0525	-2.38	1721.5	0.0175*
	F4	-0.968	374.7	0.334	0.660	823.9	0.510	-0.189	1732.3	0.850
	Cz	-0.299	378.4	0.765	1.28	823.1	0.201	2.11	1671.2	0.0352*
P200	F3	-0.209	397.2	0.835	1.14	819.5	0.253	-3.33	1734.0	0.000882***
	F4	-0.192	395.4	0.847	-1.14	803.4	0.254	-1.85	1714.3	0.0650
	Cz	0.385	397.4	0.701	-0.639	809.4	0.523	0.573	1685.9	0.567
P300	F3	-0.633	384.9	0.527	-1.28	776.8	0.202	-3.58	1713.3	0.000355***
	F4	-0.405	384.4	0.686	0.231	823.6	0.817	-2.01	1696.2	0.0442*
	Cz	-0.103	394.4	0.918	-1.08	823.8	0.281	0.0731	1667.6	0.942
N400	F3	-2.02	378.9	0.0438*	-1.83	823.9	0.0675	-2.95	1701.0	0.00327**
	F4	-1.64	371.4	0.102	0.187	823.8	0.852	-2.31	1703.9	0.0210*
	Cz	-1.45	393.5	0.148	-2.40	822.6	0.0164*	0.528	1727.8	0.597

Table 3. Detailed results of individual-level Welch's two-sample t-tests. Two-tailed tests were used. Significant effects are shown in bold, with significance levels indicated by asterisks (* $p < .05$, ** $p < .01$, *** $p < .001$).

ERP	Ch	Condition	Estimate	SE	<i>t</i>	df	<i>p</i>	Lower	Upper
N100	Cz	Intercept	-0.960	0.292	-3.30	2958	0.000988	-1.53	-0.389
		Face_happy	-0.396	0.405	-0.979	2958	0.327	-1.19	0.397
		Voice_happy	-0.248	0.407	-0.610	2958	0.542	-1.05	0.550
		Face:Voice	1.30	0.570	2.29	2958	0.0222*	0.187	2.42
	F3	Intercept	-3.29	0.512	-6.42	2958	1.55e-10	-4.29	-2.84
		Face_happy	1.70	0.711	2.39	2958	0.0171*	0.303	3.09
		Voice_happy	2.95	0.715	4.12	2958	3.94e-05***	1.54	4.35
		Face:Voice	-1.54	1.00	-1.54	2958	0.123	-3.51	0.419
	F4	Intercept	-2.59	0.467	-5.54	2958	3.33e-08	-3.50	-1.71
		Face_happy	0.670	0.649	1.03	2958	0.302	-0.603	1.94
		Voice_happy	1.80	0.653	2.76	2958	0.00586**	0.521	3.08
		Face:Voice	-0.0489	0.914	-0.0535	2958	0.957	-1.84	1.75
P200	Cz	Intercept	1.40	0.278	5.03	2958	5.20e-07	0.859	1.94
		Face_happy	-0.683	0.387	-1.76	2958	0.0773	-1.44	0.0749
		Voice_happy	-0.171	0.390	-0.441	2958	0.659	-0.934	0.591
		Face:Voice	0.171	0.544	0.313	2958	0.754	-0.897	1.24
	F3	Intercept	1.53	0.526	2.91	2958	0.00367	0.498	2.56
		Face_happy	0.935	0.731	1.28	2958	0.201	-0.499	2.37
		Voice_happy	2.03	0.736	2.75	2958	0.00591**	0.584	3.47
		Face:Voice	-2.53	1.03	-2.46	2958	0.0141*	-4.55	-0.509
	F4	Intercept	0.520	0.708	0.733	2958	0.463	-0.869	1.91
		Face_happy	0.181	0.622	0.291	2958	0.772	-1.04	1.40
		Voice_happy	1.91	0.626	3.05	2958	0.00228**	0.684	3.14
		Face:Voice	-1.94	0.876	-2.21	2958	0.0272*	-3.65	-0.218
P300	Cz	Intercept	1.77	0.329	5.38	2958	7.92e-08	1.13	2.42
		Face_happy	-0.366	0.458	-0.800	2958	0.424	-1.26	0.531
		Voice_happy	-0.112	0.460	-0.244	2958	0.807	-1.01	0.791
		Face:Voice	-0.258	0.644	-0.400	2958	0.689	-1.52	1.01
	F3	Intercept	4.70	0.658	7.16	2958	1.03e-12	3.42	5.99
		Face_happy	0.691	0.913	0.757	2958	0.449	-1.10	2.48
		Voice_happy	2.27	0.919	2.47	2958	0.0135*	0.470	4.07
		Face:Voice	-4.90	1.29	-3.81	2958	0.000140***	-7.42	-2.38
	F4	Intercept	3.08	1.62	1.90	2958	0.0577	-0.101	6.27
		Face_happy	-0.306	0.804	-0.381	2958	0.703	-1.88	1.27
		Voice_happy	1.08	0.809	1.33	2958	0.183	-0.509	2.66
		Face:Voice	-1.81	1.13	-1.60	2958	0.110	-4.03	0.408
N400	Cz	Intercept	-0.221	0.330	-0.668	2958	0.504	-0.867	0.426
		Face_happy	0.0948	0.459	0.206	2958	0.836	-0.804	0.994
		Voice_happy	0.249	0.461	0.539	2958	0.590	-0.656	1.15
		Face:Voice	-0.620	0.646	-0.961	2958	0.337	-1.89	0.645
	F3	Intercept	0.102	0.648	0.158	2958	0.875	-1.17	1.37
		Face_happy	0.508	0.901	0.564	2958	0.573	-1.26	2.27
		Voice_happy	1.80	0.906	1.98	2958	0.0476*	0.0189	3.57
		Face:Voice	-4.64	1.27	-3.66	2958	0.000253***	-7.13	-2.16
	F4	Intercept	-0.0485	1.31	-0.0369	2958	0.971	-2.62	2.53
		Face_happy	-0.344	0.801	-0.428	2958	0.668	-1.91	1.23
		Voice_happy	0.628	0.806	0.779	2958	0.436	-0.953	2.21
		Face:Voice	-2.33	1.13	-2.06	2958	0.0392*	-4.54	-0.115

Table 4. Linear mixed-effects model results for each ERP component and channel. Estimate = fixed-effect coefficient from the linear mixed-effects model; SE = standard error; Lower and Upper indicate 95% confidence intervals. Two-tailed tests were used. Significant effects are shown in bold (* $p < .05$, ** $p < .01$, *** $p < .001$).