scientific reports

Check for updates

OPEN

Rapid formation of picture-word association in cats

Saho Takagi[1,2 ✉], Hikari Koyasu[1], Miho Nagasawa[1] & Takefumi Kikusui[1]

It is well known that dogs are capable of following human verbal instructions. However, very little is known about the equivalent ability in cats. In this study, we used a switched stimuli task to examine whether cats rapidly form picture-word association, which is a fundamental ability for word learning. We presented cats with two meaningless picture-word combinations, in the habituation phase. Then, on half of the trials we switched the combination (switched condition), but the other half of the trials remained as before (non-switched condition). If cats rapidly form picture-word association, they were expected to look at the monitor for longer in the switched condition, reflecting detection of the change. We used human speech as stimuli in Exp.1, and mechanical sounds (electronic sounds) in Exp.2. Cats expressed detection of the switched combination in Exp.1, where human speech and objects were paired. However, in Exp.2 where non-social sounds and objects were paired, there was no statistical difference between switched and non-switched conditions, although there was a main effect of condition when the data from the two experiments were pooled. These results demonstrate that cats can rapidly form picture-word association. Further research should investigate whether domestication has played a role in this ability.

Humans use language to think and communicate. Understanding that a word refers to objects in the external world (referentiality) is found early in infant development. The formation of word-object associations by 8-14-month-old infants has been investigated using the procedure known as the switching task (e.g[1]). , . Infants were habituated to two word–object pairings (habituation phase) and then tested either with non-switched trials that maintained the familiar word–object pairing or switched trials consisting of the familiar word and object in a new combination. The results showed that 14-month-olds looked at the monitor for a longer duration in the switched condition after receiving 16–20 repetitions, suggesting that they rapidly formed the word-object association. Thus, human infants rapidly learned arbitrary associations between words and objects at a young age.

How did this ability to understand language evolve? It is not clear what selection pressures are related to language comprehension ability, though, some studies have shown that this word-object associative ability is not limited to humans but is present also in other species.For example, previous research has investigated language training in apes, such as the bonobo Kanzi, who demonstrated the ability to associate words with lexigrams in interactions with human caregivers[2]. African Grey Parrots are capable of comprehending auditory instructions and responding accordingly[3]. Recent research has further highlighted that companion animals, particularly those living in close proximity to humans for a long time, are also able to respond to human linguistic cues. Some dogs (called GWL: Gifted Word Learner) have been shown to possess extraordinary language comprehension abilities; for example, in one study a dog learned more than 200 object-word associations through extensive training with rewards[4]. The dog correctly fetched an object by hearing the object's name arbitrarily called by a human. Some GWL dogs can retain object-name association for at least 2 months without post-acquisition exposure[5]. Although many dogs have not been shown to exhibit object and label learning[6,7], a survey using a language inventory adapted from parent word checklists used to assess infants' language has demonstrated that domestic dogs can understand an average of 89 words[8]. This confirms that many domestic dogs are capable of following human linguistic instructions.

Recently, more attention has been paid to the social cognitive abilities of cats (*Felis catus*), the other major friend of humans. Although their ancestral form is a Libyan wild cat (*Felis Lybica*), a solitary species[9], it has been shown that modern cats read a variety of human's social cues in various modalities. For example, cats follow human pointing[10], discriminate human attentional states[11–14], refer to a human face when confronted with novelty[14,15], and discriminate human emotional expressions[16].

[1]Department of Animal Science and Biotechnology, Azabu University, 1-17-71, Fuchinobe,Chuo-ku, Kanagawa 252-5201, Japan. [2]Japan Society for the Promotion of Science, 5-3-1, Chiyoda-ku, Tokyo 102-0083, Japan. ✉email: sa-takagi@azabu-u.ac.jp

In addition to the visual cues described above, cats can also retrieve social information from auditory cues. They discriminate their owner's voice from an unfamiliar person's voice[17], mentally map the owner's position from her voice[18], predict the owner's face upon hearing her voice[19], and match emotional expressions and sounds[20]. Furthermore, a recent study showed that cats differentiate between their own name and similar-sounding nouns[21]. In addition to this sensitivity to their own name, they even represent a familiar cat's face upon hearing that cat's name[22], suggesting formation of a link between the human utterance (the cat's name) and "the cat". The study showed that cats learn a familiar conspecific's name through routine daily experience, not, explicit training involving rewards, which is also how human infants learn other people's names[22]. Thus, both species have a basic ability to quickly associate objects with speech sounds. Unlike dogs, which have been selectively bred for their ability to work with humans[23], cats have become companion animals through a process of self-domestication[24]. It is believed that only in the last few hundred years has human selection pressure been added to this process[24]. By studying cats, we can examine the relationship between self-domestication and the ability to understand language.

In this study, we tested cats on the switching task previously used with human infants[1], to examine whether cats rapidly form picture-word associations. We presented cats with two sound-picture combinations in the habituation phase, then switched the combinations on half of the trials in the test phase. If cats formed the picture-word association after brief exposure, they should detect the switch and look at the switch condition for longer. To know if the association was specific to human speech, we also conducted a control experiment in which we used electronic sounds (Exp.2). If cats have requisite capacities for vocabulary acquisition, then we might expect them to form picture-word association preferentially with human speech only.

Experiment 1
Material & methods
Ethical statement
All experimental procedures were approved by the Animal Ethics.

Committee of Azabu University (#210319-25). All methods were performed in accordance with the relevant guidelines and regulations. Informed consent was obtained from all owners. We followed the ARRIVE guidelines.

Subjects
We tested 31 cats (20 males, 11 females). Twenty-three (16 males, 7 females, mean age 3.45 years, $SD = 2.35$ years) lived in three "cat cafés", where visitors can interact and play with the cats. The other eight (4 males, 4 females, mean age 5.22 years, $SD = 3.45$ years) were household cats. We did not ask the owner to change water or food provisions. The sample size was set according to my previous similar studies of cats' social cognition experiments[19,22].

Stimuli
For each subject, we used auditory stimuli consisting of the voice of the owner calling "parumo" or "keraru", which are meaningless words. To catch the cats' attention more effectively, we asked the owner to call each word exaggeratedly and to stress the difference between them as best they could with intonation. We recorded the voices in WAV format using a handheld digital audio recorder (SONY ICD-UX560F, Japan). The sampling rate was 44,100 Hz and the sampling resolution was 16-bit. The call lasted about 0.6 s, depending on the owner (mean duration = 0.63 s, $SD = 0.13$). All sound files were normalized to a maximum amplitude of -1.0 dB using Audacity® recording and editing software, version 2.3.0. For visual stimuli, we used two pictures, one containing the "Sun" and the other containing "Pegasus". One picture was red, and the other was blue, discriminable to cats, which have dichromatic color vision. To catch the cats' attention, the pictures shrank and expanded, with the movements controlled by Python (version 3.0.0.0).

Procedure
We tested all cats individually in a familiar room. The cat was gently restrained by the experimenter 30 cm in front of the laptop computer (SurfacePro6, Microsoft) which controlled the auditory and visual stimuli. If the cat is less active, we placed the laptop in front of the resting cat at a distance of around 30 cm. Each cat was tested in one session consisting of two phases. The habituation phase consisted of 4–8 trials in which the cats were presented with word-picture combinations until each cat became habituated to the combination. One trial was defined as the playback of one of the words four times, separated by a 1 s inter-stimulus interval, with one of the pictures shrinking and expanding on the 13.3 monitor of the laptop for 9 s. All cats were exposed to one auditory stimulus at least 8 times. The word-picture combination was counterbalanced across subjects. This phase continued until the cat's looking time on the first trial was visually judged by the experimenter to have decreased by around 50%. An equal number of trials with the two word-picture combinations was presented in pseudo-random order with the restriction that the same vocalization was not repeated on consecutive trials. The inter-trials interval was about 2 s.

After each cat met the criterion, the test phase proceeded. The test phase consisted of a total of 4 trials. Each trial was the same as the habituation phase except for the switched word-picture combination in half of the trials. The inter-phase interval was about 3 min.

While Experimenter 1 gently restrained the cat, she looked down at its head; she never looked at the monitor, and so was unaware of the test condition. When the cat was calm and oriented toward the monitor, Experimenter 1 started the habituation phase by pressing a key on the computer and continuing to restrain the cat during the test. Experimenter 1 estimated looking time and made habituation judgments based on whether the cat's head was facing the monitor.
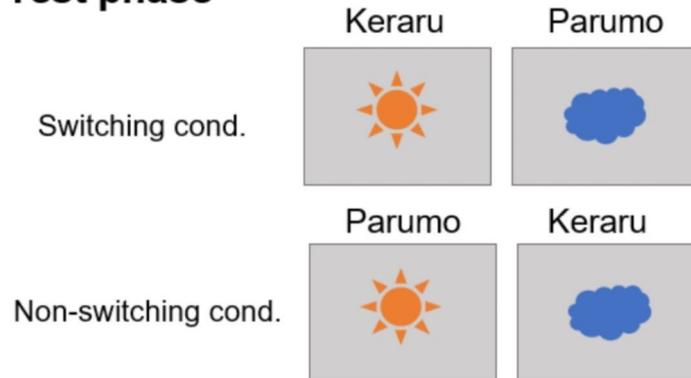
Habituation phase
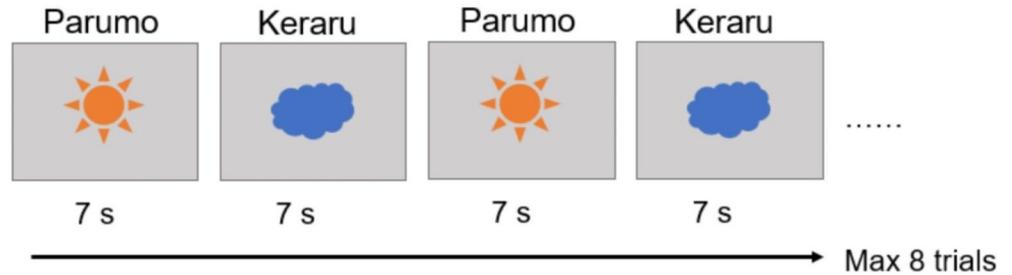
Test phase

Fig. 1. Schematic diagram of the Experiment. In the habituation phase, one of the pictures was presented on the monitor with one of the meaningless sounds. This phase proceeded until cats' reached criterion. In the test phase, there were 2 conditions involving the word-picture combination, the switched task where the combination switched and the non-switched task where the combination was the same as in the habituation phase.

Each cat's behavior was recorded by three cameras (two Gopros (HERO black 7) and SONY FDR-X3000): one beside the monitor for a lateral view, one in front of the cat to measure time looking at the monitor, and one recording the entire trial from behind.

Analysis
Individuals who looked the monitor for at least two trials during the habituation phase were included in the analysis. Data from one cat who never looked at the monitor during the habituation phase were not used. Trials in which the subject did not look at the monitor in the test phase were excluded from the analyses, whereas in the habituation phase not looking was scored as "0" as it was an important measure of habituation. A total of 31 clips were analyzed for trials 1–4. For trials 5–8, 9 clips were analyzed for trials 5 and 6, and 1 clip each for trials 7 and 8.For the test phase, 46 non-switched trials and 49 switched trials were analyzed (29 trials excluded overall). A coder who was blind to the conditions counted the number of frames (30 frames/sec.) in which the cat attended to the monitor using Adobe Premier Pro 2023. To check inter-observer reliability, N.F., also was blind to the conditions, coded a randomly chosen 20% of the videos. The correlation between the two coders' scores was high and positive (Pearson's $r = 0.96$, $n = 47$, $p < .001$).

We used R version 3.5.1 for all statistical analyses[25]. Time spent looking at the monitor was analyzed by a Generalized linear mixed model (GLMM) and a linear mixed model (LMM) using a glmer and a lmer function in a lme4 package version 1.1.10[26]. We used log-transformed looking time to get close to normal distribution in the test phase, whereas we used a poisson distribution for data in the habituation phase because it included 0 data. To examine whether cats decreased their looking time across trials, we entered trial order as fixed factor, with subject identity as a random factor in the habituation phase. To see how looking differed between conditions, condition (non-switched/ switched), living environment (café/house), and their interaction, were entered as fixed factors, with subject identity as a random factor in the test phase. We ran Wald X^2 tests using an Anova function in a car package[27] to test whether effects of each factor were significant (Fig. 1).

Results & discussion
Habituation phase
Figure 2 shows the time spent looking at the monitor during habituation phase across trials. Overall, the cats decreased their amount of time looking at the monitor across trials. GLMM revealed a significant main effect of trial order ($X^2_7 = 1395.1$, $p < .001$), suggesting that cats habituated to the stimuli.
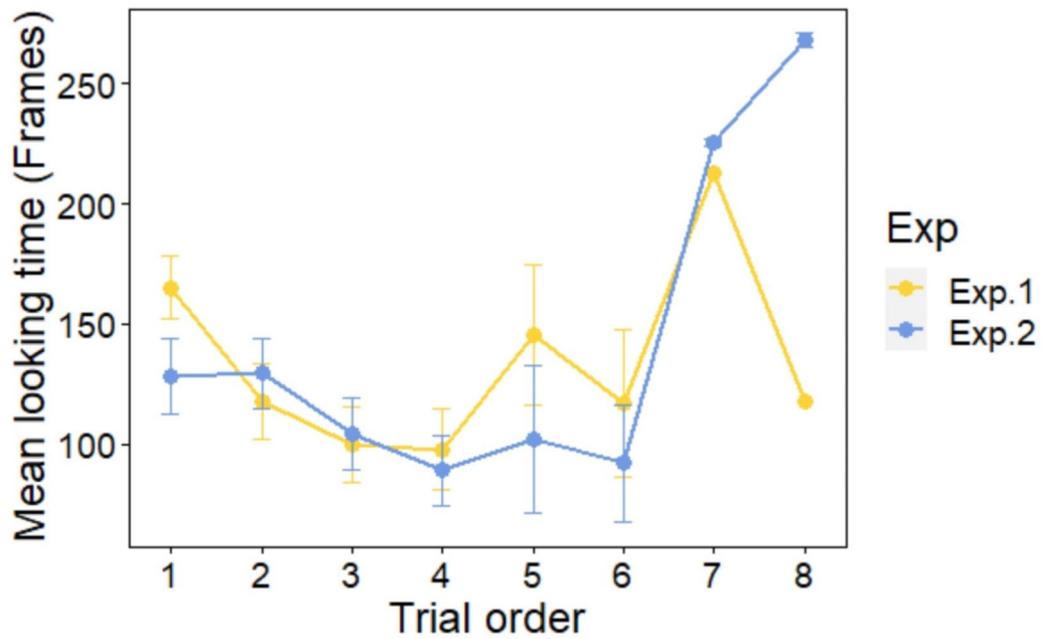
Fig. 2. Time spent looking at the monitor during habituation phase in Exp.1 and Exp.2 across trials. Error bars represent the standard error (SE). Note that there are no SEs for the 7th and 8th trials in Exp.1 as they only include one subject.
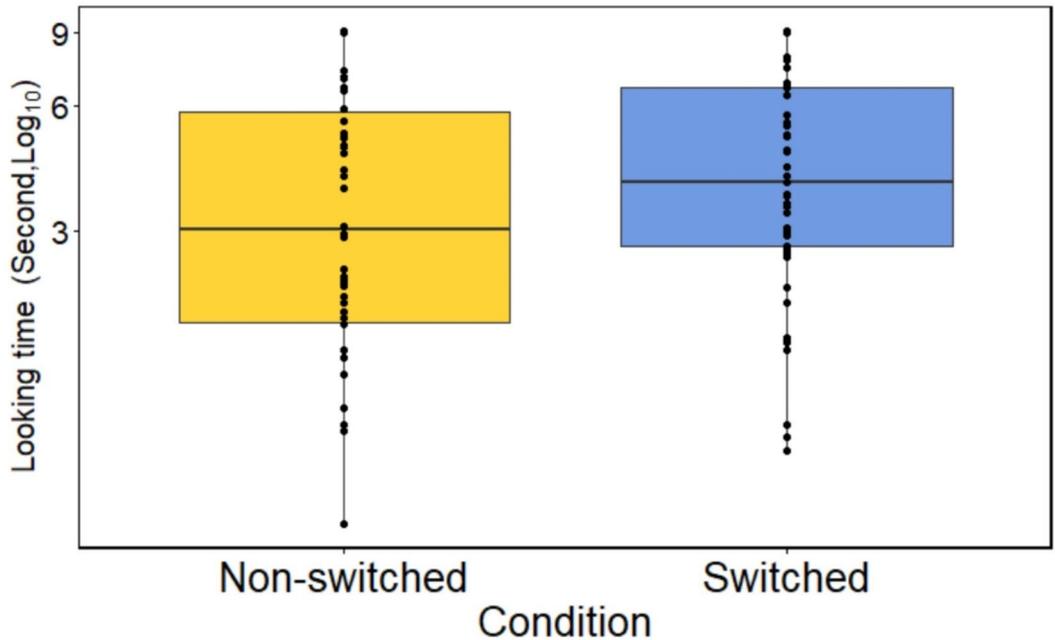

Fig. 3. Time spent looking at the monitor during test phase in Exp.1. Y axis was log-transformed to get closer to a normal distribution.

Test phase
shows time spent looking at the monitor during the test phase. Cats looked at the monitor for longer in the switched condition. LMM revealed a significant main effect of condition ($X^2_1 = 4.22$, $p = .03$). Neither the main effect of living environment ($X^2_1 = 0.94$, $p = .33$) nor the interaction ($X^2_1 = 0.28$, $p = .59$) was significant. These results indicate that cats detected the switching of the combination of pictures and sounds by forming word-picture associations. In Experiment 2, we used "physical sounds" instead of human speech to examine whether this effect was limited to a situation involving a "human" factor (Fig. 3).

4

Experiment 2

The procedure in Exp.2 was almost the same as in Exp.1, but we used non-social sounds instead of human speech.

Material & methods

Subjects

We tested 34 cats (20 males and 14 females). Twenty-four (13 males and 11 females, mean age 5.40 years, SD = 2.95 years) lived in three "cat cafés", where visitors can interact and play with the cats. The other 10 (7 males and 3 females, mean age 3.97 years, SD = 2.96 years) were household cats. Nine individuals participated in both Experiment 1 and Experiment 2. The experimental interval was at least 10 months.

Stimuli

The experimental stimuli were the same as Exp.1 except that instead of speech sounds, we used six electronic sounds downloaded from a web site that we used in a previous study[18]. All sound files were adjusted to the same volume with Audacity(R). The visual stimuli were the same as Exp.1, the sun and the Pegasus pictures. Two sound stimuli were randomly chosen for each cat.

Procedure

The procedure was the same as in Exp.1.

Analysis

We conducted almost the same statistical analysis as in Exp. 1. We measured the duration of looking at the monitor as in Exp.1. One cat escaped and climbed out of reach on the fourth trial, so this subject was only tested until the third trial in the test condition. Another cat growled on the fourth trial, so the experiment was stopped and the data up to the third trial was used. In the habituation phase, the number of clips analyzed for trials 1–8 were 34, 34, 34, 34, 11, 11, 2 and 2, respectively. In the test phase, 49 non-switched trials and 46 switched trials were analyzed (32 trials excluded overall). We analyzed the data with a GLMM / a LMM using a glmer / lmer function in a lme4 package version 1.1.10[26]. As the Exp.1, order was a fixed factor with a poisson distribution for data in the habituation phase and condition (non-switched/switched), living environment (café/house) and interactions were entered as fixed factors, with subject identity as a random factor for data both in the habituation and test phase.

Furthermore, to assess whether attention differed between Exp.1 and Exp.2, we compared looking time between experiments. For data in the habituation phase, we entered experiment (Exp.1/Exp.2) as a fixed factor, subject identity as a random factor with a poisson distribution. For the test phase, we directly compared log-transformed looking time in the test phase of Exps.1 and 2, using a LMM. We entered experiment (Exp.1/Exp.2), condition (switched/non switched), and their interaction as fixed factors and subject identity as a random factor. We ran Wald X^2 tests using an Anova function in a car package[26] to test whether effects of each factor were significant.

To check inter-observer reliability, author S.T., blind to the conditions, coded a randomly chosen 20% of the videos. The correlation between the two coders was high and positive (Pearson's $r = 0.94$, $n = 51$, $p < .001$).

Results & discussion

Habituation phase

Figure 2 shows the time spent looking at the monitor during habituation phaseacross trials. Almost all cats gradually decreased their looking time across trials. GLMM revealed a significant main effect of trial order ($X^2_7 = 769.98$, $p < .001$), suggesting that cats were habituated to the stimuli.

Test phase

Figure 4 shows that cats looked at the monitor slightly more in the switched condition, but LMM revealed no significant main effect of condition ($X^2_1 = 1.36$, $p = .24$), nor of living environment ($X^2(1) = 0.20$, $p = .65$) the interaction of condition x living environment ($X^2(1) = 0.85$, $p = .35$). These results suggest that cats did not detect the switch between pictures and electronic sounds.

Comparison between Exp.1 and Exp.2 during habituation phase

Figure 2 shows time spent looking at the monitor during habituation phase in Exp.1 and Exp.2 across trials. GLMM revealed that no significant effect of experiment ($X^2(1) = 3.40$, $p = .06$). This meant that cats were equally attentive in the habituation phase of each experiment.

Comparison between Exp.1 and Exp.2 during test phase

LMM revealed a significant main effect of condition ($X^2(1) = 5.44$, $p = .01$), but no significant effect of experiment ($X^2(1) = 0.38$, $p = .53$) or interaction between experiment and condition ($X^2(1) = 0.16$, $p = .68$). This indicated that cats formed picture-word associations regardless of sound type (human speech or an electronic sound).

General discussion

In this study, we used a switching task to examine whether cats rapidly form picture-word associations. We presented cats with two picture-wordcombinations in the habituation phase, then we switched the combination in half of the trials (switched condition) in the test phase. We predicted that if cats rapidly formed an association between sound and picture, they would detect the switch and that consequently their looking time would be
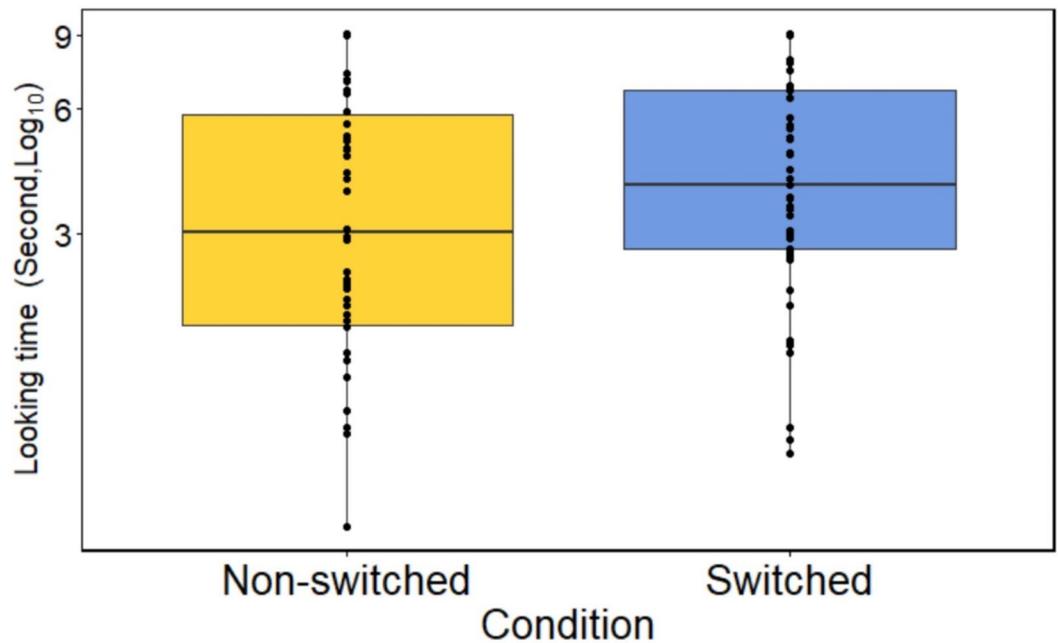
Fig. 4. Time spent looking at the monitor during test phase in Exp.2. Y axis was log-transformed to get closer to a normal distribution.

Trails in habituation pharse	Experiment 1	Experiment 2
4 trails	22	23
6 trails	8	9
8 trails	1	2

Table 1. The number of subjects of trials in habituation phase for each experiment.

longer. In Exp.1, cats looked at the monitor for longer in the switched condition, as predicted; however, this effect disappeared in Exp.2 in which we used electronic sounds instead of human speech as auditory stimuli. Although there was no interaction between experiments and conditions, we found an effect of condition, which indicated that cats were more likely to associate human speech with the picture, but this was not specific to human speech. These results suggest that cats rapidly form picture-word associations with only brief exposure, regardless of sound type.

It is noteworthy that cats made the picture-word association after only brief exposure. Most cats habituated to the stimulus pairing after 4 trials (Table 1), which means that they received only 9 s exposures in two trials for each picture-word pair. In a study of human infant, infants received at least four 20-sec trials for an picture-word pair[1]. Our results reveal that cats make associations with even less exposure. It is not clear from this study why cats can form associations very quickly. To explore this question, it should be examined from both evolutionary and developmental perspectives. This can be achieved by comparing closely related species and conducting comparative experiments on cats that are not kept by humans.

While our prediction was that cats preferentially formed the association with human speech, there were no significant interaction between Exps. and condition by a direct comparison. These results indicate that cats exhibited a tendency to focus on the switched condition in both Exp.1 and Exp.2. There was no difference between Exp. 1 and Exp. 2 to the extent that there was an interaction. However, in Exp. 2, there was no significant difference between the switched and non-switched conditions. This suggests that cats might find it easier to form associations between objects and human speech than with electronic sounds.

We used different sounds in two experiments. Although the cats appeared to pay slightly more attention to the monitor in the habituation phase in Exp.1, there was no significant difference in the total numbers of trials in the habituation phase (Table 1), which showed that attention to the monitor was similar regardless of sound type.

We did not find any behavioral difference between café and house cats in this study. Some previous studies reported differences, especially in experiments involving the owner and a stranger. For example, in one study[19] café cats showed expectancy violation whereas house cats did not, probably because of sensitivity to stranger stimuli in a visuo-auditory cross-modal task. Another study[28] also reported that café cats and house cats behaved differently when they witnessed their owner interacting with a stuffed cat. These discrepancies may stem from disparate upbringing experiences: café cats often encounter numerous unfamiliar individuals daily and may lack a consistent primary owner. One plausible explanation for the absence of differences observed in our study could

	Training		Test	
	First	Last	Non-switched	Switched
Mean	7.08	4.63	3.96	4.54
SE	0.66	0.69	0.49	0.44

Table 2. Mean looking time (sec.) and SD in Exp.1

	Training		Test	
	First	Last	Non-switched	Switched
Mean	4.67	2.68	3.9	4.14
SE	0.46	0.36	0.42	0.4

Table 3. Mean looking time (sec.) and SD in Exp.2.

be that the association-forming abilities under examination represent fundamental skills that are unaffected by experiential factors. However, the sample size of cats kept in house was smaller than that of cats in cat cafés. Future research should explore this possibility further by testing a larger number of house cats.

Cats recognize familiar cats' names without explicit training[22]. To associate the names of the familiar individuals and their referents (the name refers to "the" individual) in daily life, cats must first recognize what the human is paying attention to, and then remember what happened next (i.e., human pays attention to cat B and called cat B's name, then cat B approaches the human and sometimes gets petted or a treat). Previous studies have shown that cats recognize human attention[11–14]. The ability demonstrated in this study might be advantageous for recognizing names of other cats.

Recently, there has been a notable distinction observed in the communication styles between cats interacting with other cats and cats interacting with humans. For instance, cats employ visual cues such as gaze when communicating with humans, and they exhibit slow eye blinking as a sign of familiarity towards individuals[29,30]. This is noteworthy given that direct face-to-face communication typically indicates aggressive intentions in animal interactions[31], a pattern that may extend to cats[11]. Additionally, studies have indicated the significance of ear positions in cats for predicting subsequent cat-cat communication, with tail-up displays being rare among cats but common when approaching humans[32]. While some have characterized vocal communication in adult cats as primarily linked to conflicts or mating behaviors[33], there is evidence that cats also use vocalizations in affiliative interactions with humans[33]. It is not commonly believed that cats use vocalizations referentially with other cats. The capacity for rapid word-object mapping observed in our studies may have been developed through cats' prolonged cohabitation with humans. However, the exact origins of this ability, whether it is genetically predetermined or acquired through experiential learning, remain ambiguous. Further insights into this matter could be gained through research involving cat groups less accustomed to human interaction (Table. 2).

Some studies of dogs have examined word acquisition by reward-based training, using the fetch paradigm. Gifted dogs learn word-object associations with very few exposures[7], whereas it is difficult for many normal canines to do so[5,6]. In this study, however, we showed that "normal" cats rapidly formed picture-word associations. The main difference between our study and dog studies is in explicitness of the association. We focused on the implicit behavior of looking at the monitor, because few cats fetch an object upon command by the owner (but see[34]). Gaze has lower explicitness than fetching a toy, as developmental psychologists have pointed out (e.g[35]). , . In general, implicit behaviors such as gaze emerge before explicit behaviors. Whether cats form picture-wordassociations in situations where the focus is on explicit behaviors awaits clarification.

One limitation of our study lies in its exclusive focus on short-term associations. The inquiry into whether cats establish enduring associations, such as recalling the names of familiar cats or family members and the duration of such retention, remains unexplored. In the context of human children acquiring word meanings, the establishment of associations typically involves repeated exposure over time, resulting in storage in long-term memory[36,37]. While some gifted dogs have demonstrated the ability to exclusively select a novel toy upon hearing a novel word[4], their incapacity to recall the name of the specifically chosen toy suggests a failure in the transfer of the association to long-term memory. Notably, in cases where such transfer has occurred from experiencing object name many times, gifted dogs have retained the memory for a minimum of 2 months without further exposure[5]. The consideration of long-term perspectives in the realm of picture-word associations remains a critical avenue for future research (Table. 3).

Another limitation is that picture-word association may have been attenuated due to the fixed order of sound presentation during the habituation phase. Cats are known to rapidly habituate to stimuli in experimental settings, leading to a swift decline in interest. For instance, a study investigated cross-modal recognition of the owner's voice and face in dogs[38], using a stimulus presentation time of 30 s. In contrast[19], , who conducted a similar experiment, reduced the presentation time to 7 s. Given the propensity of cats to lose interest quickly, we employed a fixed procedure to ensure that they would focus on both stimulus sets with equal attention. However, this fixed order may have inadvertently suppressed picture-word association in the cats. Although the sound stimuli were randomized in both switched and non-switched trials during the test phase, minimizing the potential impact on the interpretation of the results, randomizing the stimuli from the habituation phase might have further strengthened the picture-word associations in the cats.

Generally, it is expected that looking time would gradually decrease from the first to the final trial of the habituation phase, as observed in infant studies. However, in our data, some cats looked longer on the final trial compared to the first trial of the habituation phase. If some individuals were not fully habituated during the habutiation phase, we would anticipate no differences between conditions in the test phase. Implementing stricter criteria for habituation may lead to clearer results. One interpretation of the rapid picture-word associations observed in this study is that the learning of these complex stimulus combinations may occur independently of word learning. The ability to learn complex stimulus combinations has been documented in many animal species (e.g.,), and this capacity in cats would be one example. Another possible interpretation is that these associations may contribute to word learning. As previously mentioned, cats are capable of learning the names of other individuals in the same household without explicit training. It is conceivable that the picture-word associations observed in this study could play a role in this process. To determine which interpretation is accurate, further studies involving additional species are necessary. Until now, word acquisition ability has been extensively explored in dogs, employing the fetch paradigm. Nevertheless, given that only a limited number of species exhibit the behavior of fetching objects for humans, the applicability of this paradigm across a broad range of species is challenging. The experimental task employed in this study exhibits adaptability for application across various species. The inquiry into whether the rapid learning to complex stimuli confirmed in this study is the ability unrelated to domestication, an outcome of evolutionary processes (especially self-domestication), or acquired through cohabitation with humans since birth remains an open question. The utilization of a comparable task across diverse species could offer insights into the evolutionary origins of this learning abilities. Investigating the universality or specificity of such capabilities has the potential to contribute significantly to our understanding of the cognitive and communicative dimensions across different species.

Data availability
I uploaded our data as supplementary materials.

References
1. Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M. & Stager, C. L. Acquisition of word-object associations by 14-month-old infants. *Dev. Psychol.* **34** (6), 1289–1309 (1998).
2. Savage-Rumbaugh, S., McDonald, K., Sevcik, R. A., Hopkins, W. D. & Rubert, E. Spontaneous symbol acquisition and communicative use by pygmy chimpanzees (Pan paniscus). *J. Exp. Psychol. Gen.* **115** (3), 211–235 (1986).
3. Pepperberg, I. M. Cognitive and communicative abilities of Grey parrots. *Appl. Anim. Behav. Sci.* **100** (1), 77–86 (2006).
4. Kaminski, J., Call, J. & Fischer, J. Word learning in a domestic dog: evidence for 'fast mapping'. *Science.* **304** (5677), 1682–1683 (2004).
5. Dror, S., Miklósi, Á., Sommese, A., Temesi, A. & Fugazza, C. Acquisition and long-term memory of object names in a sample of Gifted Word Learner dogs. *R Soc. Open. Sci.* **8** (10), 210976 (2021).
6. Fugazza, C., Dror, S., Sommese, A., Temesi, A. & Miklósi, Á. Word learning dogs (Canis familiaris) provide an animal model for studying exceptional performance. *Sci. Rep.* **11** (1), 14070 (2021).
7. Fugazza, C. et al. Rapid learning of object names in dogs. *Sci. Rep.* **11** (1), 2222 (2021).
8. Reeve, C. & Jacques, S. Responses to spoken words by domestic dogs: a new instrument for use with dog owners. *Appl. Anim. Behav. Sci.***246**, 105513 (2022).
9. Driscoll, C. A. et al. The Near Eastern origin of cat domestication. *Science.* **317** (5837), 519–523 (2007).
10. Miklósi, A., Pongrácz, P., Lakatos, G., Topál, J. & Csányi, V. A comparative study of the use of visual communicative signals in interactions between dogs (Canis familiaris) and humans and cats (Felis catus) and humans. *J. Comp. Psychol. Wash. DC 1983.* **119** (2), 179–186 (2005).
11. Koyasu, H. & Nagasawa, M. Recognition of directed-gaze from humans in cats. *Jpn J. Anim. Psychol.* **69** (1), 27–34 (2019).
12. Ito, Y., Watanabe, A., Takagi, S., Arahori, M. & Saito, A. Cats beg for Food from the human who looks at and calls to them: ability to understand humans' Attentional States. *Psychologia.* **59** (2–3), 112–120 (2016).
13. Vitale, K. R. & Udell, M. A. R. The quality of being sociable: the influence of human attentional state, population, and human familiarity on domestic cat sociability. *Behav. Processes.* **158**, 11–17 (2019).
14. Zhang, L., Needham, K. B., Juma, S., Si, X. & Martin, F. Feline communication strategies when presented with an unsolvable task: the attentional state of the person matters. *Anim. Cogn.* **24** (5), 1109–1119 (2021).
15. Merola, I., Lazzaroni, M., Marshall-Pescini, S. & Prato-Previde, E. Social referencing and cat-human communication. *Anim. Cogn.* **18** (3), 639–648 (2015).
16. Galvan, M. & Vonk, J. Man's other best friend: domestic cats (F. Silvestris Catus) and their discrimination of human emotion cues. *Anim. Cogn.* **19** (1), 193–205 (2016).
17. Saito, A. & Shinozuka, K. Vocal recognition of owners by domestic cats (Felis catus). *Anim. Cogn.* **16** (4), 685–690 (2013).
18. Takagi, S. et al. Socio-spatial cognition in cats: mentally mapping owner's location from voice. *Plos One.* **16** (11), e0257611 (2021).
19. Takagi, S. et al. Cats match voice and face: cross-modal representation of humans in cats (Felis catus). *Anim. Cogn.* **22** (5), 901–906 (2019).
20. Quaranta, A., d'Ingeo, S., Amoruso, R. & Siniscalchi, M. Emotion recognition in cats. *Animals.* **10** (7), 1107 (2020).
21. Saito, A., Shinozuka, K., Ito, Y. & Hasegawa, T. Domestic cats (Felis catus) discriminate their names from other words. *Sci. Rep.* **9** (1), 5394 (2019).
22. Takagi, S. et al. Cats learn the names of their friend cats in their daily lives. *Sci. Rep.* **12** (1), 6155 (2022).
23. Morey, D. F. The early evolution of the domestic dog. *Am. Sci.* **82** (4), 336–347 (1994).
24. Driscoll, C. A., Clutton-Brock, J., Kitchener, A. C. & O'Brien, S. J. The taming of the cat. *Sci. Am.* **300** (6), 68–75 (2009).
25. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).
26. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models using lme4 [Internet]. arXiv; [cited 2024 Jan 29]. (2014). http://arxiv.org/abs/1406.5823
27. Fox, J., Weisberg, S. & An, R. Companion to Applied Regression, Third edition. Sage, Thousand Oaks CA. [Internet]. ; (2019). https://socialsciences.mcmaster.ca/jfox/Books/Companion/
28. Bucher, B., Arahori, M., Chijiwa, H., Takagi, S. & Fujita, K. Domestic cats' reactions to their owner and an unknown individual petting a potential rival. *Pet. Behav. Sci.* **9**, 16–33 (2020).

29. Koyasu, H., Kikusui, T., Takagi, S. & Nagasawa, M. The gaze communications between dogs/cats and humans: recent research review and future directions. *Front. Psychol.* **11**, 613512 (2020).
30. Humphrey, T., Proops, L., Forman, J., Spooner, R. & McComb, K. The role of cat eye narrowing movements in cat–human communication. *Sci. Rep.* **10** (1), 16503 (2020).
31. Gallup, G. G., Cummings, W. H. & Nash, R. F. The experimenter as an independent variable in studies of animal hypnosis in chickens (Gallus gallus). *Anim. Behav.* **20** (1), 166–169 (1972).
32. Deputte, B. L., Jumelet, E., Gilbert, C. & Titeux, E. Heads and tails: an analysis of visual signals in cats, Felis catus. *Anim. Open. Access. J. MDPI.* **11** (9), 2752 (2021).
33. Bradshaw John, Cameron-Beaumont Charlotte. The signalling repertoire of the domestic cat and its undomesticated relatives. In: The domestic cat: The biology of its behaviour. 2nd edition. Cambridge University Press. pp. 67–93. (2000).
34. Forman, J., Renner, E. & Leavens, D. A. Fetching felines: a survey of cat owners on the diversity of cat (Felis catus) fetching behaviour. *Sci. Rep.* **13** (1), 20456 (2023).
35. De Bruin, L. C. & Newen, A. The developmental paradox of false belief understanding: a dual-system solution. *Synthese.* **191** (3), 297–320 (2014).
36. Mather, E. & Plunkett, K. Learning words over time: the role of stimulus repetition in mutual exclusivity. *Infancy.* **14** (1), 60–76 (2009).
37. Horst, J. S. & Samuelson, L. K. Fast Mapping but Poor Retention by 24-Month-Old infants. *Infancy off J. Int. Soc. Infant Stud.* **13** (2), 128–157 (2008).
38. Adachi, I., Kuwahata, H. & Fujita, K. Dogs recall their owner's face upon hearing the owner's voice. *Anim. Cogn.* **10** (1), 17–21 (2007).

Author contributions
S.T. participated in the design of the study, collected and analyzed the data, and drafted the manuscript as corresponding author. H.K collected data with S.T. M.N., and T.K., gave S.T. suggestions about experimental design, data analysis, discussions and helped draft the manuscript. All authors critically revised the report, commented on drafts of the manuscript, and approved the final report.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-74006-2.

Correspondence and requests for materials should be addressed to S.T.

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.