

LANGUAGE DISCRIMINATION BY LARGE-BILLED CROWS

SABRINA SCHALZ^{*1}, and EI-ICHI IZAWA²

^{*}Corresponding Author: SS3903@live.mdx.ac.uk

¹Department of Psychology, Middlesex University, London, UK

²Department of Psychology, Keio University, Tokyo, Japan

Language discrimination has previously been found in human infants, cotton-top tamarin monkeys, rats, and Java sparrows. This ability might also be relevant for the crow, a social passerine with extensive auditory perceptual skills living in close contact with humans. In this experiment we tested whether crows autonomously pay attention to spoken language, and whether they can discriminate a familiar, locally spoken language (Japanese) from an unfamiliar language (Dutch) without training. When presented with sentences spoken by multiple speakers, the crows showed significantly more responses to the Dutch than to the Japanese, which suggests that they discriminate two languages with distinctive linguistic features, and that they might also be more attentive to an unfamiliar language, Dutch, compared to a familiar one, Japanese. These results further extend the hypothesis that language discrimination is based on a general perceptual mechanism that predates the evolution of language and show that crows can voluntarily apply this mechanism to language outside of experimental set-ups.

1. Introduction

Although language as a whole is unique to humans, some cognitive abilities necessary for language are shared with other species. Ramus and colleagues (2000) showed that both human newborns and cotton-top tamarin monkeys (*Saguinus oedipus*) can discriminate between two languages from different rhythmic classes. They used Japanese (mora-timed) and Dutch (stress-timed) sentences spoken by four different female speakers per language and presented them to the infants and monkeys in a habituation/dishabituation design. In a second experiment, they synthesized these sentences to only include prosodic characteristics and removed lexical and phonetic information, as well as speaker variability (see Ramus & Mehler, 1999 for full description). They found that human infants failed to discriminate the natural stimuli, but successfully discriminated the synthesized stimuli containing only prosodic information. On the other hand, the tamarin monkeys were able to discriminate both types of stimuli despite speaker variability, although they performed better with the natural sentences than with the synthesized sentences. When presented with the synthesized stimuli played backwards, both the tamarins and the infants failed to discriminate the stimuli sets. The authors conclude from these observations that the ability to extract and process cues relevant for language discrimination likely preceded human speech, although humans and tamarins may use different cues for this task.

Taking into account a conceptual replication by Toro, Trobalon and Sebastián-Gallés (2003), this perceptual mechanism may date back even further. Their work with Long-Evans rats (*Rattus norvegicus*) using lever-press training with the original stimuli by Ramus et al. (2000) showed that just like human newborns and tamarin monkeys, rats are able to discriminate between synthesized stimuli of different rhythmical classes (stress-timed Dutch and mora-timed Japanese) when they are played forwards, but not backwards. Further research by Toro, Trobalon and Sebastián-Gallés (2005) showed that the rats were able to generalize previously learned prosodic cues to novel stimuli, and that they could only discriminate natural sentences produced by a single speaker, but not those produced by multiple speakers. It is indeed curious that all three species tested performed equally well with the synthesized stimuli, while there seems to be large variations in their performance with natural stimuli, possibly due to irrelevant information introduced by speaker variability.

The great number of parallels between birdsong and human language make passerines a well-suited model organism for biolinguistics (see Doupe & Kuhl, 1999), and their sensitivity to acoustic features can be extended to human language as well. To name just a few examples, Java sparrows can discriminate between prosodic patterns in Japanese spoken either with admiration or suspicion and generalize them to new sentences if the prosody remains familiar (Naoi, Watanabe, Maekawa & Hibiya, 2012), and discriminate English and Chinese sentences spoken by a bilingual speaker and generalize this discrimination to new sentences and a new speaker with training (Watanabe, Yamamoto & Uozumi, 2006). Zebra finches can discriminate between familiar and novel infant-directed songs and speech in English and Russian (Philmore, Fisk, Falk & Tsang, 2017), discriminate between trochees and iambs (Spierings, Hubert & ten Cate, 2017), use formant frequencies to discriminate the words wit and wet despite speaker variability (Ohms et al. 2009), and abstract prosodic patterns of human speech with prosodic stress on either the first or final syllable and generalize them to new stimuli (Spierings & ten Cate, 2014). Spierings and ten Cate (2014) conclude from this that “the sensitivity to prosodic cues is not linked to the possession of language and might have preceded language evolution, possibly originating from a pre-existing sensitivity to meaningful variation in pre-linguistic communicative sounds.”

Crows live in social groups or fission-fusion societies (Clayton & Emery, 2007), which requires them to vocally communicate with group members and identify conspecifics based on auditory cues. They can discriminate conspecifics based on their unique vocal signature (Kondo, Izawa & Watanabe, 2010), discriminate reliable and unreliable conspecifics based on their individual call (Wascher, Hillemann, Canestrari & Baglione, 2015), and recognize group members using audio-visual cues (Kondo, Izawa & Watanabe, 2012). In addition to conspecific calls, crows also discriminate between familiar and unfamiliar human voices,

possibly because they often live in close contact with humans (Wascher, Szapl, Boeckle & Wilkinson, 2012).

Taking into account these extensive capabilities related to the auditory discrimination of individual calls and the self-motivated attention to voices of both conspecifics and heterospecifics, crows may also be attentive to linguistic features of languages spoken in their surroundings. The purpose of this experiment is therefore to examine whether crows autonomously pay sufficient attention to spoken language to discriminate a familiar from an unfamiliar language without prior training. Foregoing training and keeping the experimental set-up as naturalistic as possible has the advantage of showing more accurately the linguistic capabilities and the degree of attention to linguistic features wild urban crows living in close contact with humans display on their own. We used the same stimuli previously used in Ramus et al. (2000), and Toro et al. (2003, 2005) to allow for a more accurate comparison between the findings of this experiment and the previous language discrimination experiments with human infants, cotton-top tamarin monkeys, and rats. Such a comparison might highlight the analogies and heterogeneities between these species, and thereby provide further insights into the evolution of the mechanisms necessary for language discrimination.

2. Method

Eight large-billed crows (*Corvus macrorhynchos*; four males and four females) between the ages of two and four years were tested. One female crow was excluded from analysis due to lack of response. All subjects were caught in the prefectures Tokyo, Chiba, and Ibaragi with the permission from the Environmental Bureau of the Tokyo Metropolitan Government. The crows were housed in individual stainless steel-mesh home cages with a total of twenty-four crows in the room of the animal experimental facility at Keio University. Both caretakers and previous experimenters were native Japanese speakers. Before and after the experiment, they had access to food and water ad libitum.

The experiment was carried out in an outdoor aviary (W1.5 × D2.7 × H1.6 m). In the aviary, four perches were installed in the back, middle, front and the front-right corner approximately 1m above ground. A water basin was placed on the ground. Outside the aviary, a wireless loudspeaker (Sound Link Mini, Bose, USA) was placed next to the front-right corner for stimulus presentation, and a video camera (Handy-Cam HDR CX535, Sony, Japan) for recording the crows' behaviour was placed at 50 cm from the front end of the aviary.

We used twenty Dutch and twenty Japanese sentences as stimuli. They were all declarative, adult-directed, approximately 2.5 seconds long, and spoken by four female native speakers. After the habituation to the aviary on three consecutive days, the crows were tested for their responses to the Dutch and Japanese stimuli in a total of eight trials which were distributed over four days (i.e., two trials per day). Four crows received Dutch stimuli for the first four trials and Japanese

stimuli for the last four trials, while the other three crows were assigned the opposite language order. Before the start of each trial, the crows were given 3–5 min for familiarization to the surroundings. Each trial consisted of four blocks of stimulus presentation with inter-block intervals of a 1–2-min silent period. Within each block, a set of ten sentences spoken by two different speakers was continuously presented twice in a random order. A 30 min silent period was inserted between the trials each day. The trial schedule including stimulus presentation was controlled by the programme PsychoPy 3 (Peirce, 2007). The sound level was set at a range between 70 and 80 dB across the perches. According to the different behavioural responses to 1,000 Hz and 1,600 Hz tone stimuli between individual crows in a pilot experiment, either of two behaviours as response for each crow was measured during the stimulus presentation blocks from the video-recorded data: the amount of time they had their head lifted at least above the horizontal line, or the amount of time they sat on the right half of the front perch or on the perch in the front right corner close to the loudspeaker. Response times were coded in BORIS (Friard & Gamba, 2016). To normalize the response time to the stimuli varying slightly in their durations for each crow, we calculated the relative value of response time per 10 seconds to the total stimulus duration in each (see equation 1).

$$1) \text{ relative response time} = \frac{\text{total response time}}{\text{total stimulus duration}} \times 10$$

The results were analysed using a generalized linear mixed model with an inverse Gaussian error distribution and a log link function. The model included the relative response time as an independent variable, the language as a fixed effect, and the individuals and the blocks within each trial as random effects. These analyses were performed using the free software R v.3.6.1 with the ‘lme4’ package (Bates, Maechler, Bolker & Walker, 2015). Significance of the independent variable was tested based on the Wald tests at the 0.05 level. Animal housing and the experimental protocols adhered to the guidelines of the Animal Care and Use Committee of Keio University.

3. Results

The model analysis produced a significant effect of the language variable with a negative coefficient for Japanese ($p < 0.001$, $t = -4.90$, $\beta \pm \text{S.E.} = -0.39 \pm 0.08$; figure 1). This result suggests that the crows were significantly more attentive to the Dutch sentences than the Japanese sentences.

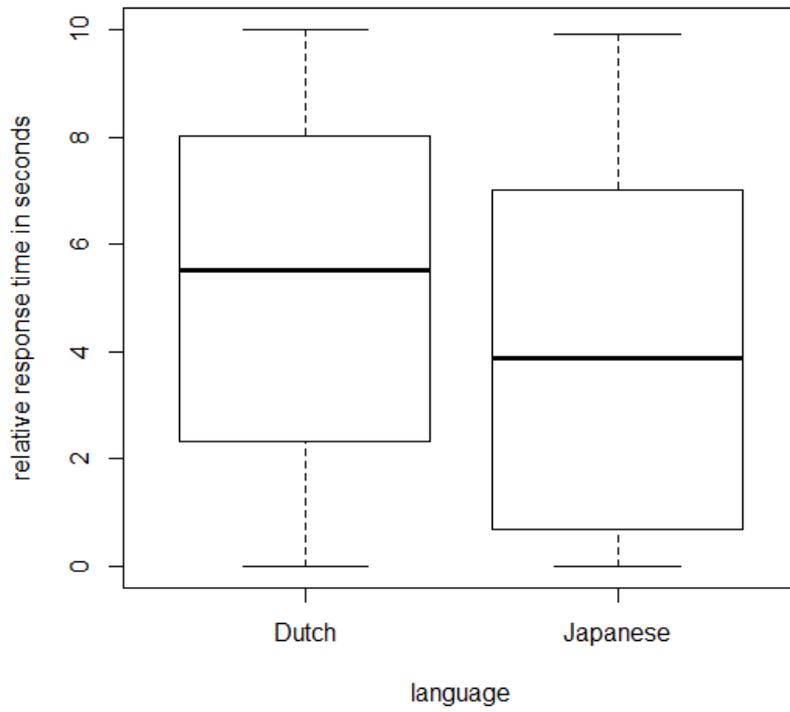


Figure 1. Relative response time per 10 seconds during the Dutch and the Japanese stimulus blocks. The crows showed more responses to the Dutch sentences than to Japanese ones.

At the individual level, five out of the seven crows clearly showed more responses to the Dutch stimuli than to the Japanese one (figure 2).

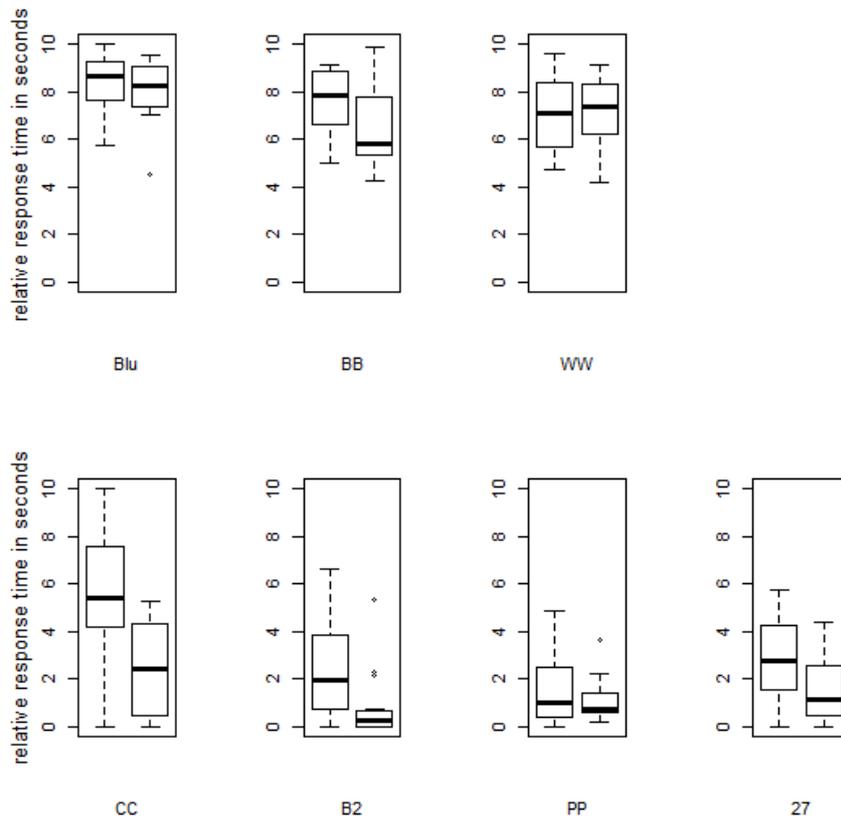


Figure 2. Individual relative response time per 10 seconds of each crow to the Dutch (left) and the Japanese (right) stimuli. The response behaviour for the crows in the first row was the amount of time their head was raised above the horizontal line, the response behaviour for the crows in the second row was the time they sat in the area next to the speaker.

4. Discussion

The results show that crows can discriminate between Dutch and Japanese sentences despite speaker variability and without prior training. A priori, crows should not be more interested in one language over the other. The initially higher attention to Dutch suggests that the crows were already familiar with Japanese before the experiment. Since all of them were caught in highly urbanized areas in Japan in and around Tokyo and were then in contact with Japanese experimenters and/or caretakers on a daily basis, it is safe to assume that they were exposed to Japanese for their entire lives. This would support the hypothesis that crows actively listen to human speech of their own accord and without experimental set-ups to a degree that would enable them to identify and later recognize key features

of Japanese independent of the individual speaker that distinguish it from other languages. Dutch, on the other hand, would likely be completely new to them and thus prompt them to pay more attention to it at first. This reaction would then be expected to gradually decline as they habituate more and more to it, and eventually their attention to Dutch should be equal to their attention to Japanese. The individual differences between the crows may be partially due to experience. “WW” and “Blu”, who were almost equally attentive to the two stimuli sets, are also the youngest crows at two and three years respectively and are considered juvenile, while the other crows are four years old. The shorter exposure to Japanese due to their young age might be the reason for their failure to discriminate it from Dutch, although further research is needed to verify this hypothesis.

Crows in urban areas such as Tokyo live in close contact with humans and speech would therefore be relevant to them, as it conveys information about the speaker’s identity and helps them determine whether they already know the specific person (Wascher et al. 2012) and whether that person might pose a threat. The perceptual abilities required for their extensive repertoire of vocalizations to communicate with conspecifics (Conner, 1985) and to discriminate group members based on their vocal signature (Kondo et al. 2010) may also be extended to the perception and categorization of human speech. Further experiments with crows from urban areas in other countries as well as crows from uninhabited areas are necessary to see whether the increased attention to the non-local language, or rather any language for crows from uninhabited areas, is consistently present.

These results stand in clear contrast with those obtained from human infants, who failed to discriminate the natural Dutch and Japanese sentences prior to the removal of non-prosodic information (Ramus et al. 2000). Speaker variability is likely the reason for this, as the rats successfully discriminated natural sentences spoken by only one speaker but failed when they were spoken by different speakers (Toro et al. 2005). The crows’ as well as the tamarins’ successful discrimination despite speaker variability points towards a more robust extraction of relevant linguistic features disregarding irrelevant information than that displayed by human infants and rats. Bird song and the vocalizations of New World monkeys show several similarities (see Snowdon, 1989), such as the repertoire of chirps and whistles used by cotton-top tamarin monkeys to convey different messages (Cleveland & Snowdon, 1982). Toro et al. (2005) argued that their experience with this type of vocalization, experience that rats do not have and infants have yet to gain, facilitates the discrimination task for the tamarin monkeys, which might also be the case with the crows. These results further support the previous findings in mammals and passerines that language discrimination is not a uniquely human ability and is instead based on a general perceptual mechanism that evolved prior to human language.

Taken together, the results obtained in this experiment show that crows living in close contact with humans are sufficiently attentive to spoken language out of intrinsic motivation to extract and recognize linguistic features distinguishing different languages from each other despite variation introduced by speaker variability. The crows' self-motivated attention to language could point towards an adaptation to sharing their habitat with humans, as an increased attention to human vocalizations might provide information on danger, comparable to eavesdropping on heterospecific alarm calls observed in multiple species (e.g. Meise, Franks & Bro-Jørgensen, 2018). This attention to linguistic features may not be limited to language discrimination or the recognition of familiar voices. Further experiments are necessary to see which elements of language animals living in urban areas are also sensitive to, and whether there are any differences compared to individuals from rural areas.

Acknowledgements

We are very grateful to Frank Ramus for kindly supplying the original stimuli used in Ramus et al. (2000, 2005) and Toro et al. (2003, 2005). This experiment was carried out as part of the Thesis@Keio programme at Keio University. The study was financially supported by JSPS KAKENHI #17H02653 and Keio University Grant-in-Aid for Innovative Collaborative Research Projects #MKJ1905 to E.-I. I.

References

- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Clayton, N. S., & Emery, N. J. (2007). The social life of corvids. *Current Biology* 17(16), R652–656.
- Cleveland, J., & Snowdon, C.T. (1982). The complex vocal repertoire of the adult cotton-top Tamarin (*Saguinus oedipus oedipus*). *Ethology* 58(3), 231–270.
- Conner, R. N. (1985). Vocalizations of common ravens in Virginia. *The Condor* 87(3), 379–388.
- Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience* 22, 567–631.
- Friard, O., & Gamba, M. (2016). BORIS: A free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7(11).
- Kondo, N., Izawa, E.I., & Watanabe, S. (2010). Perceptual mechanism for vocal individual recognition in jungle crows (*Corvus macrorhynchos*): contact call signature and discrimination. *Behaviour* 147(8), 1051–1072.

- Kondo, N., Izawa, E.I., & Watanabe, S. (2012). Crows cross-modally recognize group members but not non-group members. *Proceedings of the Royal Society B: Biological Sciences* 279(1735), 1937–1942.
- Naoi, N., Watanabe, S., Maekawa, K., & Hibiya, J. (2012). Prosody discrimination by songbirds (*Padda oryzivora*). *PloS ONE* 7(10), e47446.
- Meise, K., Franks, D.W., & Bro-Jørgensen, J. (2018). Multiple adaptive and non-adaptive processes determine responsiveness to heterospecific alarm calls in African savannah herbivores. *Proceedings of the Royal Society B: Biological Sciences* 285(1882), 20172676.
- Ohms, V. R., Gill, A., van Heijningen, C. A. A., Beckers, G. J. L., & ten Cate, C. (2009). Zebra finches exhibit speaker-independent phonetic perception of human speech. *Proceedings of the Royal Society B: Biological Sciences* (1684), 1003–1009.
- Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of Neuroscience Methods* 162(1-2), 8–13.
- Philmore, L. S., Fisk, J., Falk, S., & Tsang, C. D. (2017). Songbirds as objective listeners: Zebra finches (*Taeniopygia guttata*) can discriminate infant-directed song and speech in two languages. *International Journal of Comparative Psychology* 30, uclapsych_ijcp_32722.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288(5464), 349–351.
- Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America* 105, 512–521.
- Snowdon, C.T. (1989). Vocal communication in New World monkeys. *Journal of Human Evolution* 18(7), 611–633.
- Spierings, M., Hubert, J., & ten Cate, C. (2017). Selective auditory grouping by zebra finches: testing the iambic-trochaic law. *Animal Cognition* 20(4), 665–675.
- Spierings, M. J., & ten Cate, C. (2014). Zebra finches are sensitive to prosodic features of human speech. *Proceedings of the Royal Society B: Biological Sciences* 281(1787), 20140480.
- Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2003). The use of prosodic cues in language discrimination tasks by rats. *Animal Cognition* 6(2), 131–136.
- Toro, J. M., Trobalon, J. B., & Sebastián-Gallés, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. *Journal of Experimental Psychology. Animal Behaviour Processes* 31(1), 95–100.
- Wascher, C. A. F., Hillemann, F., Canestrari, D., & Baglione, V. (2015). Carrion crows learn to discriminate between calls of reliable and unreliable conspecifics. *Animal Cognition* 18(5), 1181–1185.

- Wascher, C. A. F., Szipl, G., Boeckle, M., & Wilkinson, A. (2012). You sound familiar: Carrion crows can differentiate between the calls of known and unknown heterospecifics. *Animal Cognition* 15(5), 1015–1019.
- Watanabe, S., Yamamoto, E., & Uozumi, M. (2006). Language discrimination by Java sparrows. *Behavioural Processes* 73(1), 114–116.