Identification of Conspecific Calls by Budgerigars (*Melopsittacus undulatus*)

Camille Toarmino
University at Buffalo, State University of New York
Erikson G. Neilans
University at Buffalo, State University of New York
Micheal L. Dent
University at Buffalo, State University of New York

Humans have evolved recognition systems that allow them to extract information from signals, such as words, that are not fully intact. The present study explored whether budgerigars may have evolved a similar recognition system that allows for the identification of conspecifics (i.e., members of the same species) even when their calls are not fully audible. We hypothesized that the accuracy of the budgerigars' categorizations would increase when 1) more vocal material was presented and 2) when the beginnings, rather than the ends, were present. Using operant conditioning techniques, four budgerigars were trained to identify two different contact calls. Each bird was then presented with the task of categorizing different pieces of those calls, and recognition accuracy was measured. As predicted, we found that the amount of vocal material and the presence of the beginning of calls both facilitate recognition accuracy. Specifically, we found that the beginning of a call facilitates recognition when at least 50% of a call is available. Implications for budgerigar and human vocal recognition are discussed.

The recognition of mates, prey, kin, or territory is crucial for the survival of many species. The recognition of conspecifics (i.e., members of the same species) may especially increase the reproductive fitness of individual organisms within a species (Ward & Schlossberg, 2004; Rabenold, 1984). In humans, vocal recognition of other people facilitates effective communication by allowing them to identify senders or receivers of communicative signals when other recognition cues are unavailable (e.g., Garrido et al., 2009). The recognition of conspecifics is beneficial among budgerigars (*Melopsittacus undulatus*), or small Australian parrots. Budgerigars are a colonial species, with individual organisms immersed in a cacophony of conspecific calling and environmental noise (Brockway, 1964). In such noisy environments, budgerigars must be able to vocally identify mates or kin in order to increase reproductive success (Lengagne et al., 1999). Humans are also able to identify another person's voice in noisy settings (the cocktail party phenomenon; Cherry, 1953). It would be advantageous for budgerigars to have a comparable ability, because colonial settings often result in acoustic information being interrupted or completely obscured. Humans have evolved recognition systems that allow them to extract information from signals, such as words, that are not fully intact (Salasoo & Pisoni, 1985). The present study explored whether budgerigars may have evolved a similar recognition system that allows for the identification of conspecifics even when their calls are not fully audible.

Budgerigars produce a variety of vocalizations, such as alarm calls and contact calls. Alarm calls may function to alert the group of a potential threat (Brockway, 1964), whereas contact calls are emitted typically when budgerigars are separated (Brockway, 1964) and may function to identify conspecifics (Park & Dooling, 1985). Budgerigars must be able to discriminate between alarm calls and contact calls in order to make use of the unique information encoded in each type of call. For instance, correctly identifying alarm calls in the presence of predators may increase the chance of survival for individuals within the group. However, if an alarm call were to become obscured before it is perceived, then its specific information pertaining to a threat may be lost. Therefore, it seems especially advantageous for budgerigars and similar species to have evolved recognition systems that can still extract meaningful information from obscured signals.

There is evidence to suggest that some species do not require entire vocalizations for the recognition of conspecifics. European starlings (*Sturnus vulgaris*) can recognize conspecifics significantly above chance after hearing only 1-2 seconds of a conspecific song (Knudsen, Thompson, & Gentner, 2010). Furthermore, increasing the length of a vocalization also increases recognition accuracy (Knudsen, Thompson, & Gentner, 2010). The black-legged kittiwake (*Rissa tridactyla*) is another species that can recognize conspecifics without hearing entire vocalizations. Black-legged kittiwake chicks respond more vigorously to calls of their parents even when less than the entire call is presented, suggesting that they are able to identify obscured parental calls (Mulard, Aubin, White, Hatch, & Danchin, 2008). Budgerigars are similar to starlings and black-legged kittiwake gulls in that they are all highly vocal and colonial species. Because all three species tend to live in noisy, colonial environments, it would be beneficial for them to have evolved recognition systems that do not need full vocalizations for identification. We therefore expected that budgerigars, similar to starling and black-legged kittiwake gulls, would be able to recognize conspecifics without requiring entire vocalizations. We also expected that increasing the length of their vocalizations would increase recognition accuracy.

However, not all portions of a vocalization may be equally helpful in facilitating recognition. In humans, the beginnings of words are most effective in facilitating recognition compared to other portions of words (Desroches et al., 2009; Marslen-Wilson et al., 1989; Salasoo & Pisoni, 1985). Salasoo and Pisoni (1985) presented human participants with several sentences, of which one word was partially occluded by noise. The words were occluded from either the beginning or the end by consecutive 50 millisecond increments on each successive trial. Participants needed to hear only about 50% of the occluded word in order to identify it in a meaningful context. For example, in the sentence "The stray cat gave birth to kittens," participants need to hear only half of the word "kittens" in order to correctly identify it. Importantly, participants were better able to identify words when the ends, rather than beginnings, of words were occluded. This may be because the beginning of a word contains acoustic-phonetic information that aids identification. As more of a word is presented, it may increasingly narrow down an originally large word candidate pool, eliminating incorrect word choices and providing more opportunity for correct word identification.

The beginning of a signal may not be facilitative of recognition in all species. Knudsen et al. (2010) found that no effect was observed for the testing of initial, middle, or terminal parts of a European starling song in a recognition task. In other words, no single part of the song was more reliable than the others in inducing recognition. However, we expected to find that for the budgerigar, the beginning would be the most salient piece of a call for inducing recognition. Because starling songs are much longer than budgerigar calls, starlings have ample material available to recognize an individual even if part of the song is obscured. In contrast, budgerigar calls are milliseconds in duration and provide little opportunity to recognize a caller if a portion of the call is obscured. It would be beneficial for the beginning of a call to aid in recognition because it may help to immediately narrow down call candidates and thus facilitate recognition. Therefore, in budgerigars, the beginning of a call may be an ideal location for cues that lead to conspecific identification.

Beginning the recognition process almost as immediately as a signal is received seems especially important for highly vocal species such as budgerigars and humans. Understanding the recognition system of budgerigars may provide insight into how the recognition system of humans has evolved, changed, and been preserved over time. In addition, examining an animal with a similar recognition system to humans may yield understanding for disorders such as phonagnosia, the impaired ability to identify familiar voices with the people to whom they belong (Van Lancker & Canter, 1982). Furthermore, if it is the case in budgerigars, like in humans, that the beginning of a vocalization is most facilitative of recognition, this may suggest that call and word perception is strongly shaped by the acoustic environment in which a species resides.

The present study investigated whether budgerigars are able to recognize conspecifics when less than the entire contact call is available. We sought to determine the average amount of vocal material needed for accurate recognition in budgerigars, and whether material at the beginning, middle, or end of the call is particularly important for this recognition. After being trained to categorize two different contact calls, budgerigars were required to categorize pieces of the original calls. We varied the amount and location (beginning or end) of vocal material that was presented. We hypothesized that the accuracy of the budgerigars' categorizations would increase when 1) more vocal material was presented and 2) when the beginnings, rather than the ends, were present.

METHOD

Subjects. Four adult budgerigars (two male and two female) were used in this experiment. All birds were housed individually in a vivarium at a large public university and were kept on a strict day/night cycle according to the season. The birds were obtained through breeding in the laboratory or bought at a local pet store. They were tested five consecutive days a week, twice a day in 30-40 minute sessions. They were maintained at 85-90% of their free-feeding weights throughout the experiment. All procedures were approved by the university's Institutional Animal Care and Use Committee and complied with NIH guidelines for animal use.

Testing Apparatus. The birds were tested in a wire cage (61 x 33 x 36 cm) placed inside a sound-attenuated chamber (Industrial Acoustics Company, Small Animal Chamber). The test cage consisted of a perch, automatic food hopper (Med Associates Standard Pigeon Grain Hopper), and two keys that were vertically aligned in front of the bird. Each key was a microswitch with a 1-cm square button glued to the bottom end. A 7-W light illuminated the test cage. A web-camera (Logitech QuickCam Pro, Model 4000) was placed in the test cage to monitor bird activity at all times during each session. A speaker (Morel Acoustics, Model MDT-29) was placed in the test cage about 30.5 cm away from the bird during testing. Experimental files were made through Sykofizx software controlled by Dell microcomputer Tucker-Davis Technologies (TDT, Gainesville, FL) modules.

Stimuli. In order to obtain natural bird calls for our experiments, six contact calls were recorded from five individual budgerigars (three male and two female) in the laboratory. These birds were obtained from the same room where the experimental birds were housed. Therefore, these calls were familiar to the birds used in the experiments.

To obtain the recordings, each bird was placed individually into an echo-reducing box. This box was made of cardboard and lined with sound absorbing foam (10.2 cm Sonex, Ilbruck Co.). A condenser microphone (PRM902) was then placed inside the box about six inches away from the bird. Another bird was placed in the room in order to incite calling. The bird's responses were recorded and edited in Adobe Audition. Each call was 500ms in duration and split into 125ms quarters (1, 2, 3, and 4). The combined quarters were 250ms (1+2 and 3+4) and 375ms (1+2+3 and 2+3+4) in duration.

Procedure. There were two phases of this experiment: the training phase and the testing phase. Each budgerigar was trained using operant conditioning to peck keys, receiving food pellets as reinforcement. In the training phase, a budgerigar was trained to identify two different conspecific training calls that were played at equal probability. The bird was trained to associate the left key with one call and the right key with a different call. When the calls were correctly identified, the hopper was illuminated by a light and the bird was allowed access to food in the hopper for 1.5 seconds. When the calls were incorrectly identified, the hopper light was extinguished for five seconds and the bird did not receive access to food. There were a total of three different experimental sessions for each bird, each with two different contact calls. Once a bird reached 80% recognition accuracy consistently for at least 300 trials on one of the pairs of calls, it moved on to the testing phase of that session. In the testing phase, the training calls were still presented, but only on 80% of the trials. The other 20% of the trials consisted of different isolated and combined quarters of the calls. The isolated call stimuli were created by cutting the whole call into quarters and labeling them by their serial position within the call (e.g., 0-25%=1, 25-50%=2). The combined call stimuli were created by adding isolated stimuli starting with the first quarter (e.g., 1+2, 1+2+3), and without the first quarter present within the call (e.g., 0-25%=1, 25-50%=2). The combined call stimuli were created by adding isolated stimuli starting with the first quarter (e.g., 1+2, 1+2+3), and without the first quarter present (e.g., 3+4, 2+3+4). Every time an isolated quarter or combined quarter stimulus was played and a bird pecked a key, it was reinforced regardless of whether or not the response was correct. This was done to prevent a bird from associating a piece of the call with one response over another, and was the reason that these test stimuli were only presented on a small proportion of all trials. Once the birds completed 20 trials for each test stimulus, they were retrained on another pair of contact calls and tested, followed by a third training and testing session. Each bird was tested on the pairs of calls in a different order than the other birds. The percentage correct for each bird was calculated using the last 20 trials.

RESULTS

A preliminary one-way ANOVA revealed no statistically significant difference in recognition accuracy between the 2nd, 3rd, and 4th isolated quarter conditions, p=.100. These three conditions were equivalent in that the same amount of call (25% of the entire call) was presented, and the beginning of the call was absent. In subsequent analyses, we therefore averaged results from these three conditions into a single score.

We conducted a 3 (Amount of call: 25%, 50%, 75%) × 2 (Call content: Beginning present vs. beginning absent) repeated-measures ANOVA on recognition accuracy (see Figure 1). We found a main effect of Amount of call, $F(2, 6) = 6.02$, $p = .037$, $\eta_p^2 = .67$, as well as a main effect of Call content, $F(1, 3) = 33.96$, $p = .010$, $\eta_p^2 = .92$. Importantly, these main effects were qualified by a significant interaction, $F(2, 6) = 5.36$, $p = .046$, $\eta_p^2 = .64$. Analyses of simple main effects revealed that when subjects were presented with 25% of a call, there was no statistically significant difference in recognition accuracy when the beginning of the call was present ($M = 52.10$, $SD = 11.99$) or absent ($M = 47.79$, $SD = 9.20$), $p = .129$. However, when subjects were presented with 50% of a call, they demonstrated significantly higher recognition accuracy when the beginning of the call was present ($M = 63.13$, $SD = 4.78$) than when the beginning of the call was absent ($M = 52.01$, $SD = 3.34$), $F(1, 3) = 29.29$, $p = .012$, $\eta_p^2 = .91$. Similarly, when subjects were presented with 75% of a call, they demonstrated significantly higher recognition accuracy when the beginning of the call was present ($M = 74.17$, $SD = 7.79$) than when the beginning of the call was absent ($M = 59.18$, $SD = 3.91$), $F(1, 3) = 19.43$, $p = .022$, $\eta_p^2 = .87$.

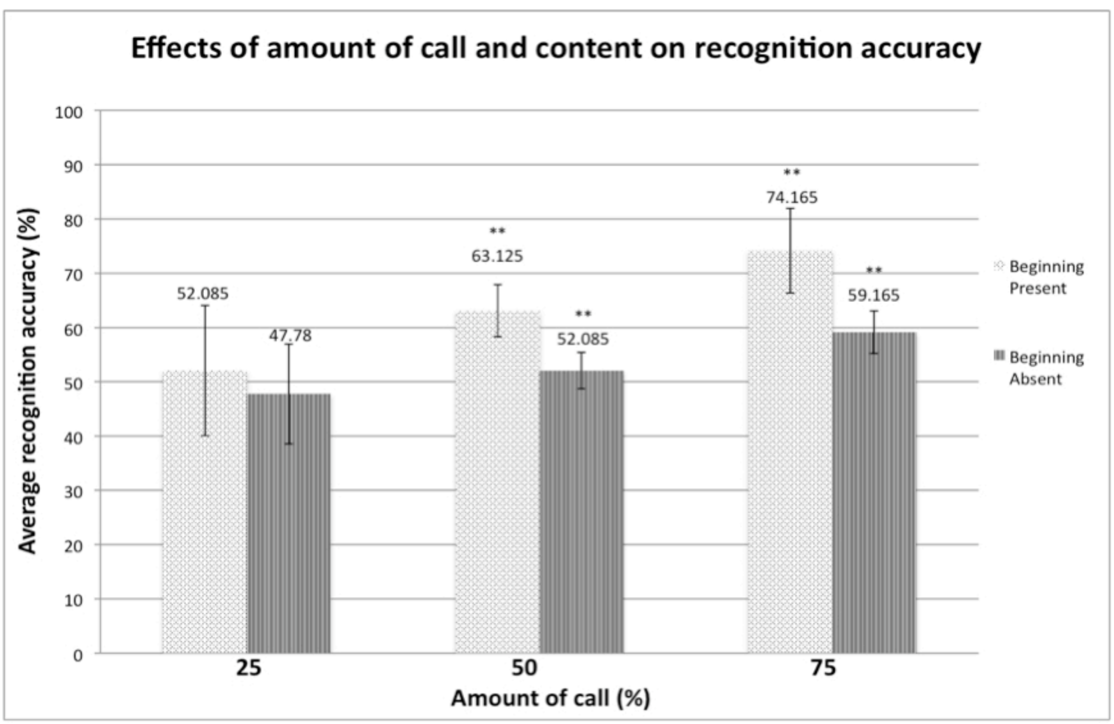

FIGURE 1. Average percent correct of experimental stimuli. Significant differences between opposing content conditions (e.g., 25% beginning present vs. 25% beginning absent) are represented by ** = $p < .05$. Error bars represent between-subjects standard deviations. Chance performance is at 50%.

DISCUSSION

The present study explored vocal recognition of conspecifics in budgerigars, as well as potential parallels between human and animal communication. Budgerigars, like starlings, black-legged kittiwake, and humans, appear to require only part of a vocalization for identification. This is a particularly advantageous feature for highly social species, which often live in boisterous environments in which vocalizations are frequently degraded and obscured. Furthermore, we found that recognition accuracy increased as more of a vocalization was presented, and that whether the beginning of the call facilitated recognition was dependent on how much of the call was presented. Specifically, we found that when only 25% of a call was presented, the beginning of the call did not significantly facilitate recognition, but that when 50% or 75% of a call was presented, the beginning of the call significantly facilitated recognition.

Similar to previous findings regarding European starling song perception (Knudsen et al., 2010), we found that budgerigars are better able to recognize conspecifics when more of a call is presented. However, unlike with European starling songs, in which the beginning of the song is no more facilitative of recognition than the middle or end of the song, we found that in budgerigar calls, the beginning of the call can facilitate recognition. This difference may be explained by the fact that the length of European starling songs differs greatly from that of budgerigar contact calls. European starling songs used in the experiment of Knudsen et al. (2010) were around 12 seconds in duration and often contained redundant motifs, whereas the budgerigar contact calls used in the present study were 500 milliseconds in duration and contained considerable variability throughout the call. Because starling songs are much longer than budgerigar calls, it is possible that starlings are attending to a distribution of cues throughout the song, whereas budgerigars are paying close attention to the beginning of these very short utterances.

This aspect of the budgerigar recognition system may be similar to the way humans perceive and retrieve words from the mental lexicon. In some models of word recognition in humans, such as the neighborhood activation model (Luce and Pisoni, 1998), one of the earliest stages deals with processing acoustic-phonetic information. It would be beneficial for the encoding of acoustic-phonetic information to begin almost as immediately as the signal is perceived, because bottom-up processing of a signal may accelerate the process of eliminating false word candidates. However, in highly vocal environments, such as those of humans and budgerigars, it is unlikely that a signal is encoded strictly from its acoustic-phonetic information because in these environments, the presence of noise often degrades acoustic-phonetic information. Therefore, a recognition system that accounts for (but is not entirely reliant on) acoustic-phonetic information, such as in the neighborhood activation model, may result in optimal signal encoding.

Furthermore, the finding that the beginnings of budgerigar calls can facilitate recognition suggests that there may be a recognition cue encoded in the beginning of a call. In this respect, call perception in budgerigars and human speech perception are very similar. In humans, the phonetic information contained in the beginning of a word aids in identification (Salasoo & Pisoni, 1985). However, it is still unclear in both budgerigar call perception and human speech perception as to whether it is the content of the beginning information or the serial position (i.e., the beginning being the first part of the signal that is encoded) of that information within a call or word that aids recognition. Because the content of budgerigar contact calls is highly variable, it is unlikely that it is strictly the content of that information that facilitates recognition. It is more likely that the serial position within the call (i.e., the beginning) facilitates recognition. If the beginning of a call were occluded, budgerigars may begin to attend to the middle (which would be the "new" beginning of the call) when trained on these "new" calls.

However, it is possible that *both* the content of the call and the serial position of information influence identification. In human speech production, the beginning syllable of a word is most likely to be stressed or "strong" (Cutler & Carter, 1987). Stress on the first syllable increases the saliency of that syllable, and subsequently changes the information content from that of the rest of the word. This finding suggests that there may be an interaction between the serial position of the beginning information and the actual information content of the beginning that allows for more accurate word identification. In other words, even if a word is somehow invariable throughout, stress on the beginning syllable changes the information content of the rest of the word, making the beginning more salient and easier to recognize. Future research may benefit from exploring the relative importance of call content and the serial position of information in the call in call identification. If information content is more important, budgerigars should still demonstrate accurate recognition if a call were manipulated to have the beginning information located elsewhere within the call. If the beginning information were eliminated, budgerigars should demonstrate significantly less accurate recognition (as seen in the current study). If serial position is more important, then budgerigars should be able to alter which parts of the call they attend to when the beginning of the call is moved.

The present research has extended previous work on human speech perception, suggesting that even species as seemingly different as humans and budgerigars may share similarities between their recognition systems. Although no animal model is a flawless representation of the human condition, the budgerigar shares certain similarities with humans concerning the cognitive processing and perception of behaviorally relevant acoustic signals. Although many experiments concerning speech and signal perception can be directly tested on humans, it is important to discover how these systems work in animal models so that more advanced testing of human communicative disorders can be done. Importantly, exploiting key similarities between humans and budgerigars, such as vocal learning, may provide insight into the evolution of language. Without the ability to modify and learn words, a key aspect of vocal learning, spoken language would be nearly impossible. Vocal learning must have evolved before spoken language. With this idea in mind, researchers can begin to explore other research topics using the budgerigar model, such as the conditions in which vocal learning arose, how social environments have shaped signal perception, and how different modes of auditory recognition have evolved.

References

Brockway, B. F. (1964). Ethological studies of the budgerigar *(Melopsittacus undulatus)*: non-reproductive behavior. *Behaviour*, 22, 193-222.

Charlton, B. D., Zhihe, Z., & Snyder, R. J. (2009). The information content of the giant panda, *Ailuropoda melanoleuca,* bleats: acoustic cues to sex, age, and size. *Animal Behaviour*, 78, 893-898.

Cutler, A., & Carter, D. M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech and Language*, 2, 133-142.

Desroches, A., Newman, R. L., & Joanisse, M. (2009). Investigating the time course of spoken word recognition: Electrophysiological evidence for the influences of phonological similarity. *Journal of Cognitive Neuroscience*, 21, 1892-1906.

Dooling, R. J. (1986). Perception of vocal signals by budgerigars (*Melopsittacus undulatus*). *Experimental Biology*, 1986, 195-218.

Dooling, R. J., Brown, S. D., Klump, G. M., & Okanoya, K. (1992). Auditory perception of conspecific and heterospecific vocalizations in birds: evidence for special processes. *Journal of Comparative Psychology*, 106, 20-28.

Dooling, R. J., Park, T. J., Brown, S. D., Okanoya, K., & Soli, S. D. (1987). Perceptual organization of acoustic stimuli by budgerigars (*Melopsittacus undulatus*): II. Vocal signals. *Journal of Comparative Psychology*, 101, 367-381.

Farabaugh, S. M., & Dooling, R. J. (1996). Acoustic communication in parrots: Laboratory and field studies of budgerigars, *Melopsittacus undulatus.* In D. E. Kroodsma & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds (*pp. 97-117).New York, Cornell University Press.

Garrido, L., Eisner, F., McGettigan, C., Stewart, L., Sauter, D., Hanley, J. R., Schewinberger, S.R., Warren, J. D., Duchaine, B. (2009). Developmental phonagnosia: A selective deficit of vocal identity recognition. *Neuropsychologia*, 47, 123-131.

Holekamp, K. E., Boydston, E. E., Szykman, M., Graham, I., Nutt, K. J., Birch, S., Piskiel, A., Singh, M. (1999). Vocal recognition in the spotted hyaena and its possible implications regarding the evolution of intelligence. *Animal Behaviour*, 58, 383-395.

Knudsen, D., Thompson, J. V., & Gentner, T. Q. (2010). Distributed recognition of natural songs by European starlings. *Learning and Motivation*, 41, 287-306.

Lengagne, T., Jouventin, P., & Aubin, T. (1999). Finding one's mate in a king penguin colony: Efficiency of acoustic communication. *Behaviour,* 136, 833-846.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1-36.

Marslen-Wilson, W., & Zwitserlook, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology*, 15, 576-585.

Mulard, H., Aubin, T., White, J. F., Hatch, S. A., & Danchin, E. (2008). Experimental evidence of vocal recognition in young and adult black-legged kittiwakes. *Animal Behaviour,* 76, 1855-1861.

Nordeen, K. W., & Nordeen, E. J. (1992). Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches. *Behavioral and Neural Biology*, 57, 58-66.

Okanoya, K., & Dooling, R. J. (1991). Perception of distance calls by budgerigars (*Melopsittacus undulatus*) and zebra finches (*Poephila guttata*): assessing species-specific advantages. *Journal of Comparative Psychology*, 105, 60-72.

Owings, D. H., & Coss, R. G. (1977). Snake mobbing by California ground squirrels adaptive variation and ontogeny. *Behaviour,* 62, 50-69.

Park, T. J., & Dooling, R. J. (1985). Perception of species-specific contact calls by budgerigars *(Melopsittacus undulatus). Journal of Comparative Psychology*, 99, 391-402.

Rabenold, K. N. (1984). Cooperative enhancement of reproductive success in tropical wren Societies. *Ecology*, 65, 871-885.

Salasoo, A., & Pisoni, D.B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language*, 24, 210-231.

Stoddard, P. K., & Beecher, M. D. (1983). Parental recognition of offspring in the cliff swallow. *The Auk* 100, 795-799.

Van Lancker, D. R., Canter, G. J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain Cognition*, 1, 185-195.

Ward, M. P., & Scholossberg, S. (2004). Conspecific attraction and the conservation of territorial songbirds. *Conservation Biology*, 18, 519-525.