3aAB9. Auditory scene analysis in budgerigars (Melopsittacus undulatus) and zebra finches (Taeniopygia guttata)

Micheal Dent*, Erikson Neilans, Mary M. Flaherty and Amanda K. Martin

*Corresponding author's address: Psychology, University at Buffalo, SUNY, B76 Park Hall, Buffalo, NY 14260, mden@buffalo.edu

Deciphering the auditory scene is a problem faced by humans and animals alike. However, when faced with overlapping sounds from multiple locations, listeners are still able to attribute the individual sound objects to their individual sound-producing sources. Here, we determined which characteristics of sounds are important for streaming versus segregating in birds. Budgerigars and zebra finches were trained using operant conditioning procedures on an identification task to peck one key when they heard a whole zebra finch song and to peck another when they heard a zebra finch song missing a middle syllable. Once the birds were trained to a criterion performance level on those endpoint stimuli, probe trials were introduced on a small proportion of all trials. The probe songs contained modifications of the incomplete training song's missing syllable. When the bird responded as if the probe was a whole song, it suggests they streamed together the altered syllable and the rest of the song. When the bird responded non-whole song, it suggests they segregated the altered probe from the rest of the song. Results show that some features, such as spectrotemporal similarity and location, are more important for streaming than other features, such as timing.

Published by the Acoustical Society of America through the American Institute of Physics
AUDITORY STREAMING IN BIRDS

Acoustic communication by animals in the real world can often be a difficult process, with many senders of many species sending many signals through noisy environments. Yet, animals and humans alike are able to attribute specific signals to one individual sound-producing object or another fairly easily. The ability to do this is known as auditory stream segregation (Bregman, 1990). Auditory streaming has been extensively studied in European starlings (Sturnus vulgaris; e.g., Hulse et al. 1997; MacDougall-Shackleton et al. 1990), but not much is known about it in other birds, and not much is known about the numerous cues important for streaming versus segregating acoustic stimuli in animals. Presentation rate and frequency separation are, of course, important for streaming, but other characteristics of acoustic signals known to be important for streaming in humans, such as location, frequency content or bandwidth, timing, and intensity have yet to be studied in animals. Here, auditory streaming of a zebra finch (Taeniopygia guttata) song was measured in zebra finches and budgerigars (Melopsittacus undulatus) using operant conditioning techniques. The importance of spatial, intensive, spectral, and temporal information on auditory streaming was characterized in several experiments, described below.

METHODS

Budgerigars and zebra finches were trained using operant conditioning experiments to peck keys for food reinforcement. The birds were trained on an identification task to first identify two endpoint stimuli (see Fig. 1). The two training endpoints were a whole zebra finch song (called “whole” in this manuscript) and a zebra finch song with one syllable deleted (called “broken” in this manuscript). The birds first pecked a left key to start a trial. After a 2-7 s variable interval, one of the two endpoints was presented from a speaker located at 90 degrees to the left of the bird (except the final experiment, which used the 0 degree speaker) with equal probability. Half of the birds were trained to peck the left key to the whole and the right key to the broken songs, while the other half was trained on the opposite conditions. If a bird correctly identified the song within 2 s of its presentation, it was reinforced with 1.5 s access to hulled millet from an illuminated standard pigeon grain hopper on 70% of all trials and an illuminated hopper for 1.5 s on the other 30% of all trials. If the bird incorrectly identified the endpoint training stimuli, the house light was extinguished for 5 s. The birds could immediately initiate another trial after the reinforcement or punishment phases were completed by pecking the left microswitch key once. After the birds were able to correctly identify the two endpoint training stimuli at a rate of at least 80% correct for 300 consecutive trials, they were moved onto the testing phase of the experiment. During the testing phases of the experiments, the training stimuli were still presented on 80% of all trials, and a correct response was still required for reinforcement. On the probe trials, however, since there was no true “correct” or “incorrect” answer, the birds were always reinforced, as long as they responded within the appropriate response window to one of the two keys. These probe trials were presented on only 20% of all trials so that the birds did not develop response biases towards one category. The first 20 trials for each probe type were analyzed and compared to the responses to the two endpoint stimuli.

The probe trials were interspersed randomly with the training trials, one experiment at a time, in a random experimental order for each subject, until all data were collected. In the first experiment, the effects of spatial location on streaming were measured. In these trials, the broken song was presented from the left (-90 deg) speaker, as it was in training. In addition, the missing syllable (which would have made it a whole song) was presented from a speaker at 0 or +90 deg. The syllable was presented at the time it would have been presented in the whole song condition. If the bird responded as hearing a whole song during the probe where the broken song plus the missing syllable from another location was presented, it suggests that the birds streamed together the information from the two locations into one auditory object. In the second experiment, the effects of intensity were measured. In these trials, the broken song was presented at one intensity and the missing syllable was presented at a lower or higher intensity than the rest of the song. If the birds responded as hearing a whole song during the probe where the missing syllable was at an intensity much lower or much higher than the rest of the song, it suggests that intensity is not much of a factor in auditory streaming. In the third experiment, the effects of timing on auditory streaming were measured. Many studies have shown that presentation rate has a strong influence on auditory streaming (reviewed in Bregman, 1990). The relative timing of objects within the stream is less well known. Here, the missing syllable was presented slightly earlier or slightly later than it would have been within a normal song. That is, if the syllable was usually presented at 50 ms after the previous syllable, here it would be presented 30 or 80 ms after the previous syllable. If the bird responded as hearing a whole song during the probe where the broken song plus the missing...
sylable at the wrong time was presented, it suggests that temporal cues within a song are not critically important for streaming. In the fourth experiment, the role of spectrotemporal cues was measured. In this experiment, the broken song was presented as usual, along with a filtered version of the missing syllable. The missing syllable was high or low-passed relative to the rest of the song. If the bird responded as hearing a whole song during the probe where the missing syllable was filtered, it suggests a relative unimportance of spectral cues during auditory streaming. Finally, in the last experiment, the content of the missing syllable was investigated. In all other experiments, the missing syllable that was presented during the probe trials was some version of the exact syllable used during training. Here, however, the missing syllable was a completely different syllable. The probe stimuli syllables were: other zebra finch syllables, zebra finch contact calls, budgerigar contact calls, and budgerigar warble elements. Rates of streaming for spectrotemporally similar versus dissimilar syllables were measured in this experiment.

FIGURE 1. Spectrograms of the zebra finch song training stimuli – the “whole” song (A) and “broken” song (B). The whole song contains five syllables while the broken song has the fourth syllable deleted and replaced with silence.

RESULTS AND CONCLUSIONS

Results show that the birds often reported as hearing a “whole” song during probe trials. This suggests that certain conditions elicited streaming. There was a range of responses across experiments and conditions, however. Spatial location, spectral similarity, and intensity were all factors that influenced streaming, while temporal fine structure had little influence on streaming. Probe trial syllables presented from a different location than the original song produced more “broken” responses than those from the same location, especially at a 180 degree spatial separation. Probe trial syllables that were less intense than the broken song produced more “broken” responses than those at the same intensity as the rest of the song, while syllables that were more intense than the rest of the song were usually streamed. Probe trial syllables presented too early or too late within the song almost always elicited a streaming response from the birds. Probe trial syllables that were high-passed were less likely to be streamed than
complete or low-passed syllables. Finally, different syllables than the original missing syllable were more likely to be streamed if they were spectrotemporally similar to that original syllable than when they were different.

Auditory streaming allows the auditory system to make use of a number of cues within auditory signals to correctly attribute them to their proper sound-producing sources. This is the first study to show that a number of the Gestalt principles of organization are important for auditory streaming in birds, allowing for similar items to be grouped together and different items to be split from one another. Here, results show that similarity in location, intensity, duration, and spectral content all allow birds to correctly stream objects. Studies such as these highlight the similarities in auditory processing and organization across animal species.

ACKNOWLEDGMENTS

Special thanks to Barbara Shinn-Cunningham and her laboratory for providing us with the stimuli, to Kristen Garcia, Kelly Radziwon, Thomas Welch, and the many undergraduates who assisted with this project. Supported in part by an NSF OBSSR Science of Learning Center Supplement.

REFERENCES