# JARO
Journal of the Association for Research in Otolaryngology

# Spectral Contrasts Underlying Auditory Stream Segregation in Goldfish (*Carassius auratus*)

RICHARD R. FAY

*Parmly Hearing Institute and Department of Psychology, Loyola University Chicago, Chicago, IL 60626 USA*

## ABSTRACT

This study investigates the effects of spectral separation of sounds on the ability of goldfish to acquire independent information about two simultaneous complex sources. Goldfish were conditioned to a complex sound made up of two sets of repeated acoustic pulses: a high-frequency pulse with a spectral envelope centered at 625 Hz, and a low-frequency pulse type centered at 240, 305, 390, or 500 Hz. The pulses were presented with each pulse type alternating with an overall pulse repetition rate of 40 pulses per second (pps), and a 20-pps rate between identical pulses. Two control groups were conditioned to the 625-Hz pulse alone, repeated at 40 and 20 pps, respectively. All groups were tested for generalization to the 625-Hz pulse repeated alone at several rates. If the two pulse types in the complex resulted in independent auditory streams, the animals were expected to generalize to the 625-Hz pulse trains as if they were repeated at 20 pps during conditioning. It was hypothesized that as the center frequency of the low-frequency pulse approached that of the 625-Hz pulse, the alternating trains would be perceived as a single auditory stream with a repetition rate of 40 pps. The group conditioned to alternating 625- and 240-Hz pulses generalized least, with maximum generalization at 20 Hz, suggesting that the animals formed at least one perceptual stream with a repetition rate of 20 pps. The other alternating pulse groups generalized to intermediate degrees. Goldfish can segregate at least one "auditory stream" from a complex mixture of sources. Segregation can be based on spectral envelope and grows more robust with growing spectral separation between the simultaneous sources. Auditory stream segregation and auditory scene analysis are shared among human listeners, European starlings, and goldfish, and may be primitive characteristics of the vertebrate sense of hearing.

Keywords: goldfish, auditory stream segregation, spectral separation

## INTRODUCTION

Conditioning experiments on complex sound perception have revealed a sense of hearing for goldfish that is qualitatively similar to that of many avian and mammalian species, including humans (Fay 1970; 1972; 1988; 1992a, b; 1994a, b; 1995; 1998a, b; Fay et al. 1996). Most recently, a form of auditory stream segregation was demonstrated for goldfish (Fay 1998a). An auditory stream is the perceptual correlate of a sound source or event that is formed by analyzing sound features from a mixture of sources arriving at the ears and then parsing them to generate useful hypotheses about the identity and characteristics of individual sources. Determining one source from within a mixture of sources is termed "auditory stream segregation" in human hearing (Bregman 1990). Auditory stream segregation is an important component of auditory scene analysis, a fundamental characteristic of the human sense of hearing (Bregman 1990).

It has been argued that auditory scene analysis may be a fundamental feature of a sense of hearing shared by most vertebrates (Fay 1998b). Conditioned behavior consistent with this notion has been obtained from

*Correspondence to:* Richard R. Fay • Parmly Hearing Institute and Department of Psychology, Loyola University Chicago • 6525 North Sheridan Road • Chicago, IL 60626. Telephone: (508) 548-6469; fax: (508) 548-2003; email: rfay@luc.edu

European starlings using natural bird song signals (Hulse et al. 1997; Wisniewski and Hulse 1997) and more abstract pure tone sequences (MacDougall–Shackelton et al. 1998), and from goldfish (Fay 1998a). Bregman (1990) has emphasized that ". . .there really are distinct physical objects and events in the world that we humans inhabit. Therefore the evidence that is obtained by our senses really ought to be untangled and assigned to one or another of them" (p. 13). It now appears that this complex perceptual process may be common among vertebrates. Most usual acoustic environments are made up of multiple and often simultaneous sound sources. We suggest that appropriate behaviors with respect to these sources would require perceptual processes analogous to stream segregation and scene analysis. Thus, it is not surprising that vertebrates, in addition to human beings, are capable of this sort of processing.

In a previous study of stream segregation in goldfish (Fay 1998a), the method of stimulus generalization with classical, respiratory conditioning was used to demonstrate that two simultaneous sources composed of pulse trains with different repetition rates (19 and 85 pps) and spectral envelopes (centered at 238 and 625 Hz, respectively) were analyzed independently. Goldfish were initially conditioned to a mixture of the two sources and then tested for response to each individual source type at a number of different repetition rates. It was found that generalization (response strength) declined toward increasing repetition rates of the 238-Hz source and toward lower repetition rates of the 625-Hz source. In other words, goldfish correctly associated the pulse type with the appropriate repetition rate following conditioning to a simple mixture of the two sources. As is the case for human listeners, it was further found that having one source begin 500 ms before the other (onset asynchrony) enhanced the stream segregation when goldfish were tested for generalization to the leading source type. In this experiment, the spectral envelopes of the sources (238 and 625 Hz) were widely spaced within the goldfish audiogram in order to maximize the possibility of stream segregation.

The present experiments investigate the effect of spectral separation on the goldfish's stream segregation behavior. Our analysis of spectral separation effects makes use of a fundamental and early demonstration of auditory stream segregation in human listeners (Miller and Heise 1950). In that case, a sequence of tones (i.e., ABABAB. . ., where A and B are different in frequency) can result in two types of perceptions depending on the frequency separation between A and B and the rate of repetition. When A and B are very close in frequency, they appear to form a single entity alternating in pitch; this was termed a "trill" by Miller and Heise (1950). At larger frequency

differences, however, the perception of the sequence breaks into two unrelated and interrupted tones, or two auditory streams that repeat at the rate of A or B presentation. This phenomenon was termed "perceptual fission" by van Noorden (1975; 1977). He also showed that fission could occur on the basis of alternating tone levels as well as frequencies. This behavior for human listeners has been modeled as a primitive process characteristic of low-level physiological processing (Beauvois and Meddis 1996) and as a more complex process that includes possible attentional mechanisms (McCabe and Denham 1977). Bregman (1990) argued that these sorts of effects are "pre-attentive" because they occur spontaneously without conscious effort.

MacDougall–Shackelton et al. (1998) have investigated tone sequence perception in European starlings. Animals were trained to discriminate between two types of temporal patterns of brief tone presentation; one, termed galloping, comprised repeating groups of three equally spaced tone pulses, and the other, termed isochronous, comprised single pulses repeated at the gallop rate or at twice the gallop rate. In stream segregation tests, probe trials were introduced that consisted of a galloping pattern with the central pulse of each three-tone group differing in frequency from the flanking pulses. Starlings responded to these probe stimuli as if they were isochronous when the frequency difference was sufficiently large. Thus, frequency contrasts caused a galloping temporal pattern to apparently split into separate streams repeating at the rates of the isochronous conditioning stimuli. In combination with previous studies of stream segregation in starlings making use of bird song stimuli (Hulse et al. 1997; Wisniewski and Hulse 1997), the authors argued that auditory scene analysis is an adaptive phenomenon that is likely to be found in other species.

As noted above, an aspect of scene analysis has now been demonstrated for goldfish using brief tone pulses differing in repetition rate and spectral envelope (Fay 1998a). One hypothesis arising from that experiment was that stream segregation in goldfish depends, as it does in human listeners (e.g., Bregman and Campbell 1971), on the spectral differences, or contrasts, between the two potential sources that occurred simultaneously. In order to investigate this hypothesis and extend the generality of the original stream segregation demonstration in goldfish, we used a new segregation test similar to the ABAB . . . alternating sequence paradigm common in studies on human listeners. In the new test, animals were conditioned to pulse trains made up of alternating pulse types (ABAB . . .); A with a high (625 Hz) center frequency, and B with a lower (240–500 Hz) center frequency. Animals were then tested for generalization to the 625-Hz pulse type repeated at a variety of rates. It was hypothesized that

if segregated auditory streams were formed based on the spectral difference between sources, goldfish would generalize to the 625-Hz pulse type as if it had been repeated at one half the overall pulse repetition rate used in conditioning (i.e., AAA... at 20 pps, or the actual repetition rate of the 625-Hz pulse in conditioning). On the other hand, if the sources were not segregated, the generalization gradient would resemble that obtained following conditioning to the 625-Hz A source alone, repeated at twice the rate (i.e., 40 pps). The effect of the magnitude of the spectral contrast would be seen as evidence for stream segregation at larger contrasts, but not for smaller contrasts.

## METHODS

### Animals

Subjects were 48 common goldfish (*Carassius auratus*) about 8 cm in length from snout to tail, maintained in aquaria for three weeks to several months. The care and use of animals was approved by the Institutional Animal Care and Use Committee of Loyola University Chicago.

### Experimental design

All experiments consisted of classical respiratory conditioning sessions followed by stimulus generalization test sessions. Eight animals were conditioned and tested in each experiment, and each animal served once in only one experiment. In general, the experiments determined whether and under what conditions goldfish behave as if the alternating pulse trains (ABAB...) were segregated into two streams, each at half the repetition rate (the rate of AAA..., or BBB..., or 20 pps) of the conditioning mixture (overall rate of 40 pps). As in Fay (1998a), the rationale for this experimental design comes from a previous experiment using similar stimuli presented singly (Fay 1995). It was found that both the repetition rate and the pulse spectral profile determined the response in generalization tests. In the previous set of experiments, conditioning to one stimulus (e.g., the 238-Hz pulse repeated at 19 pps) was followed with a generalization test in which response magnitude declined as repetition rate was varied from that used in conditioning. These generalization gradients showed that pulse repetition rate controls generalization behavior robustly and is a salient stimulus dimension for the goldfish. Control experiments 1 and 2 of the present group are an extension of the previous experiments. In control experiment 1, animals were conditioned to the 625-Hz pulse repeated at 20 pps and then tested for generalization to the same pulse repeated at a variety of rates.
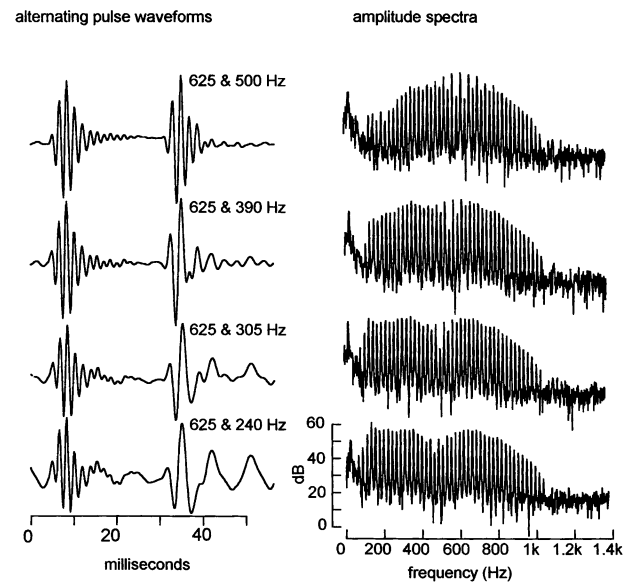


**FIG. 1.** Stimuli: **Left**. Sound pressure hydrophone recordings (5-kHz sample rate) of the alternating pulse trains used in conditioning the experimental groups, filtered between 100 and 2500 Hz. The leftmost pulse is always the 625-Hz pulse, and the following pulse is the lower-frequency, alternating pulse. The hydrophone replaced the fish in the restraint bag. **Right**. Amplitude spectra (Matlab "spectrum" FFT) of these stimuli computed using 4096 points of the waveform samples, excluding waveform onsets and offsets.

In control experiment 2, the conditioning stimulus was the 625-Hz pulse repeated at 40 pps, with the same generalization test stimuli as in control experiment 1. In experiments 3–6, the conditioning stimulus (ABAB...) was made up of the 625-Hz pulse (A) alternating with a lower-frequency pulse (B: 240, 305, 390, or 500 Hz, respectively). For these experiments, the generalization test stimuli were the same as for control experiments 1 and 2; the 625-Hz pulse repeated at 20, 23, 27, 33, 40, 49, 62, and 80 pps. Figure 1 illustrates these various pulse types as hydrophone recordings of them in the test tank (described below).

It was hypothesized that if the two pulse types of the alternating conditioning stimuli led to stream segregation, one of the pulse types alone would be perceived as having a repetition rate equal to its actual rate of repetition in the alternating series used in conditioning (one-half the overall repetition rate, or 20 pps). Thus, the generalization gradient for the control group conditioned to the 625-Hz pulse repeated at 20 pps should resemble the generalization gradient for the groups conditioned to the alternating series, if stream segregation occurred. If stream segregation did not occur, the generalization gradient for the control group conditioned to the 625-Hz pulse repeated at 40 pps should resemble the gradient for the groups conditioned to the alternating pulses. The degree of resemblance between the experimental and control

gradients would indicate the degree to which a particular alternating pulse pair led to segregated streams.

## Acoustics and stimuli

The test chamber was a Plexiglas cylinder 23 cm in diameter and 28 cm high. A swimming pool loudspeaker (University UW-30) was buried in sand at the bottom of the tank with the diaphragm facing up and about 2 cm below the sand surface. The chamber rested on a limestone slab inside an Industrial Acoustics Inc. single-walled audiometric booth.

The five pulse types used in these experiments were synthesized by multiplying a sinusoid (240, 305, 390, 500, or 625 Hz) by a 10-ms Gaussian envelope such that a single waveform peak occurred at the envelope's center. Thus, the synthesized pulses were temporally symmetric. For the control and all generalization test stimuli, trains were synthesized by repeating the 625-Hz pulses at the following rates: 20, 23, 27, 33, 40, 49, 62, and 80 pps. Conditioning stimuli for the alternating series (experiments 3–6) were created by alternating the 625-Hz pulse type with the lower-frequency pulse type with 25 ms between successive pulses. Thus, the overall repetition rate for the alternating pulses was 40 pps, while the time between identical pulses was 50 ms (equivalent to a repetition rate of 20 pps).

For all experiments, the conditioning stimulus was a 6-s presentation of a single periodic pulse train (control experiment 1 and 2), or an alternating series of two pulse types (experiments 3–6). Stimuli were presented at a level at which the frequency component at the center of the spectral profile was about 40 dB above absolute threshold for pure tones at that frequency (Fay 1969). During conditioning (described below) stimulus levels were varied randomly within a 24-dB range from trial to trial but were always set at the center of this range during generalization tests. Stimuli were synthesized and fed through a digital-to-analog converter at 5 kHz, low-pass filtered at 1500 Hz, controlled in amplitude by a programmable attenuator, amplified by a Crown 100-W power amplifier, and then led to the loudspeaker.

Acoustic signals were measured by a Bruel & Kjaer 8104 calibrated hydrophone placed in the fish restrainer (see below). The hydrophone output was amplified by a Bruel & Kjaer Nexus charge amplifier, filtered between 100 and 2500 Hz, and digitally recorded. Alternating pulse waveforms and their power spectra are shown in Fig. 1. The spectra were computed on 4096 point recorded waveforms using the Matlab "spectrum" fast Fourier transform routine. In the spectra, the overall envelopes reflect the spectra of individual pulses, and the lines reflect the repetition rate. Pulse spectra are wide due to the brief pulse durations. The roll-offs toward the lowest frequencies are truncated by the attenuating frequency response of the test tank below 100 Hz.

## Conditioning and testing

These methods were described in detail previously (e.g., Fay, 1995). Goldfish were gently restrained in a cloth bag about 2 cm below the water surface and in the center of the test chamber. A thermistor was placed near the mouth to measure water flow during respiratory mouth movements. Respiration was defined as the length of the recorded waveform in arbitrary units minus the length with no respiratory activity occurring. The respiratory response was defined as the ratio of the respiratory waveform's length during the last 4 s of the stimulus to the sum of the latter and the waveform's length 4 s preceding the stimulus [the suppression ratio (SR)]. Respiratory suppression lasts several seconds (unconditioned response) following a mild, 100-ms AC electric shock (unconditioned stimulus). Shock was delivered through steel screen electrodes placed at the animal's head and tail and at the time the conditioned stimulus ended.

Goldfish were conditioned and tested in one or two sessions. A session included 40 conditioning trials (random intertrial intervals at an average of 3 min). Generalization test trials then occurred in the same session or in another session the next day. Generalization tests consisted of 40 trials (average intertrial interval of 1.5 min) and included eight test stimuli presented four times each in random order without shock. During the generalization test, the conditioning stimulus was presented with shock every fifth trial to maintain conditioned response magnitude.

A median SR was calculated for each of the generalization stimuli presented during the test session. Generalization was normalized relative to the median SR to the conditioning stimulus measured during the generalization session and expressed as a percentage. Percentage generalization was defined as $[(0.5 - \text{med}_\text{T}) / (0.5 - \text{med}_\text{C})] \cdot 100$, where $\text{med}_\text{T}$ is the median SR to the generalization test stimulus, and $\text{med}_\text{C}$ is the median SR to the conditioning stimulus. Percent generalization values above 100% occurred in the rare cases that the respiratory suppression response to a novel test stimulus was greater than that to the conditioning stimulus.

## RESULTS

### Control experiments 1 and 2

Animals in control groups were first conditioned to trains of the 625-Hz pulse at 20 pps (experiment 1)
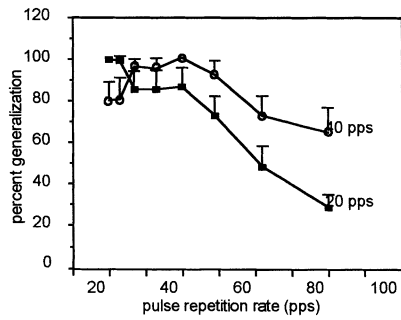
**FIG. 2.**  Group mean generalization gradients for two groups of eight animals in control experiments. Vertical bars indicate the standard error of the means. These means and 95% confidence intervals are in Table 1. For group 1 (labeled 40 pps; open circles), animals were initially conditioned to the 625-Hz pulse type repeated at 40 pps and then tested for generalization to the 625-Hz pulse repeated at the several rates indicated on the abscissa. For group 2 (labeled 20 pps; filled squares), animals were conditioned to the 625-Hz pulse type repeated at 20 pps and then tested for generalization as in group 1.

and 40 pps (experiment 2) and then tested to the 625-Hz pulse type which was repeated at eight repetition rates logarithmically spaced between 20 and 80 pps. Group generalization gradients for eight animals were obtained by averaging individual gradients. Figure 2 shows these group mean gradients. Table 1 lists these means under the 40 pps and 20 pps columns along with 95% confidence intervals about each mean.

Both generalization gradients are rather broad with substantial response strength at many repetition rates. However, generalization tends to decline nearly monotonically as the repetition rate of the test stimuli deviates from that used in conditioning. Clearly, the pulse repetition rate exerts control over generalization behavior. The differences between these two gradients are small, but all features of the gradients are consistent with the difference between the conditioning repetition rates used. The gradient for the 20-pps rate

(filled squares) has a maximum at 20 and 23 pps and a more rapid decline toward higher repetition rates than does the gradient for the 40-pps rate (open circles). The 40-pps rate gradient has a maximum at 40 pps and declines somewhat shallowly for repetition rates higher and lower than the maximum. These results resemble those of previous, similar experiments (Fay 1972, 1994a, 1995, 1998a). They show that goldfish behave as if periodic pulses evoke a perceptual dimension that has elements in common with the dimensions of roughness or periodicity pitch estimated by human listeners (Terhardt 1970). Regardless of the interpretation, the differences between these two generalization gradients and the resemblance between either of these gradients and the experimental group (3–6) gradients are the criteria that are used in the remaining experiments that investigate the effects of spectral contrasts on auditory stream segregation.

### Experiments 3, 4, 5, and 6

The four experimental groups were first conditioned to an alternating mixture of the 625-Hz pulse (A) and a lower-frequency (B) pulse in an ABAB... pattern. They were then tested for generalization to the same set of test stimuli used in control experiments 1 and 2. These group mean gradients are shown in Figure 3A. Table 1 lists these plotted means and their 95% confidence intervals. All are broad gradients, differing primarily in the overall degree of generalization and in details of their shapes. The gradient for the alternating 625- and 240-Hz pulses shows the least generalization, and its shape resembles the control gradient for animals conditioned using the 625-Hz pulse repeated at 20 pps (Fig. 2, filled squares). We interpreted the similarity in shape to mean that both group 2 (control, 20 pps) and group 3 (experimental, alternating 625 and 240 Hz) respond in generalization as if the repetition

**TABLE 1**

Percent generalization means and 95% confidence intervals about each mean for the six groups plotted in Figures 2 and 3A.

|  | 40 pps[a] | | 20 pps | | 240[b] | | 305[b] | | 390[b] | | 500[b] | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Test pps | mean | CI (+−)[c] | mean | CI (+−) | mean | CI (+−) | mean | CI (+−) | mean | CI (+−) | mean | CI (+−) |
| 20 | 80.1 | 20.9 | 100 | 0.0 | 65.4 | 29.4 | 69.2 | 13.6 | 72.6 | 14.4 | 83.2 | 20.8 |
| 23 | 80.9 | 24.1 | 99.4 | 5.5 | 67.9 | 33.1 | 74.1 | 16.9 | 82.4 | 15.6 | 90.4 | 12.0 |
| 27 | 96.7 | 8.1 | 85.5 | 29.2 | 60.0 | 31.7 | 78.0 | 15.9 | 78.2 | 15.8 | 89.9 | 11.4 |
| 33 | 95.7 | 12.2 | 85.6 | 21.4 | 61.3 | 33.1 | 79.7 | 17.3 | 88.3 | 9.2 | 84.8 | 8.2 |
| 40 | 100 | 0.0 | 86.9 | 21.6 | 64.3 | 28.1 | 75.7 | 14.0 | 92.9 | 10.1 | 102 | 8.2 |
| 49 | 92.5 | 16.3 | 72.9 | 22.2 | 40.3 | 27.6 | 63.9 | 21.9 | 68.7 | 28.2 | 102 | 11.6 |
| 62 | 72.9 | 22.6 | 48.4 | 23.6 | 28.6 | 22.8 | 52.5 | 23.2 | 61.0 | 15.5 | 90.3 | 18.4 |
| 80 | 64.8 | 27.6 | 28.7 | 14.9 | 23.1 | 17.9 | 45.8 | 19.0 | 48.4 | 19.5 | 71.7 | 27.3 |

[a] apps = pulses per second.
[b] 240–500: Center frequencies (Hz) of pulses alternating with 625-Hz pulses in conditioning.
[c] CI(+−): 95% confidence intervals about each mean.
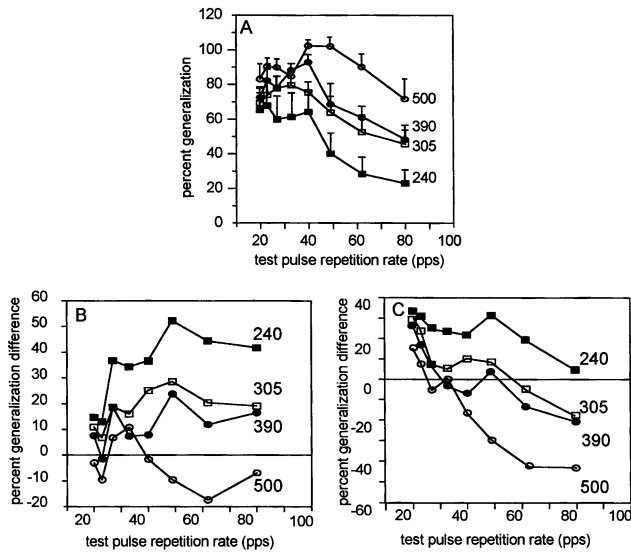
**FIG. 3.** **A**. Group mean generalization gradients for experimental groups 3–6. Animals were conditioned to alternating pulses [the 625-Hz pulse type alternating with a lower-frequency pulse (240, 305, 390, or 500 Hz pulse types, respectively)] and then tested for generalization as in groups 1 and 2. Vertical bars indicate the standard error of the means. Means and 95% confidence intervals appear in Table 1 **B**. Generalization gradients in **A** were subtracted from the control group 1 gradient (Fig. 2, open circles) to give the percent generalization difference between the response of animals conditioned to a single pulse train at 40 pps (control group 1) and those conditioned to a mixture of pulses alternating at the same overall repetition rate (experimental groups 3–6). Differences are largest for the greatest spectral contrast between alternating pulses and smallest for least spectral contrast. **C**. Generalization gradients in **A** were subtracted from the control group 2 gradient (Fig. 2, filled squares) to give the percent generalization difference between the response of animals conditioned to a single pulse train at 20 pps (control group 2) and those conditioned to a mixture of pulses alternating at the same overall repetition rate (experimental groups 3–6). Functions are flattest for the greatest spectral contrast between alternating pulses and steepest (pps-dependent) for least spectral contrast.

rate were 20 pps, indicating some degree of stream segregation. The overall lower generalization for group 3 (625 Hz alternating with 240-Hz pulses) is thought to arise from the essential dissimilarity between the alternating pulses (conditioned stimulus) and the nonalternating pulses of only one type (generalization test stimuli) for group 3. Thus, even though the animals of group 3 may have segregated two streams, each of which repeated at 20 pps, these two streams are certainly not equivalent to the single stream evoked during generalization testing. This essential dissimilarity helps to explain the relatively small amount of generalization, overall, for group 3.

The gradient for the alternating 635- and 500-Hz pulses shows the most generalization, and its shape resembled the control gradient for animals conditioned using the 625-Hz pulse repeated at 40 pps (group 1, Fig. 2, open circles). The shape similarity

suggests that both group 1 and group 6 respond in generalization as if the conditioning repetition rate was 40 pps, indicating that the two sources for group 6 (625 Hz alternating with 500-Hz pulses) were essentially fused to evoke a single stream. This interpretation helps explain the overall greater amount of generalization exhibited by group 6. The other two generalization gradients (groups 4 and 5) are intermediate between groups 3 and 6 in amount of generalization and shape, indicating intermediate degrees of stream segregation.

In Figure 3B, each experimental group mean gradient is subtracted from the gradient of control group 1 (625-Hz pulse, 40 pps), providing an indication of the similarity between these respective gradients. Experimental group 3 (625 Hz alternating with 240-Hz pulses in conditioning) clearly is the least similar, differing by as much as 50 percentage points at high repetition rates. Figure 3B also illustrates that as the alternating pulse types become more similar in spectral envelope, the gradients monotonically approach that of control group 1 (625 Hz, 40 pps; 0% generalization difference). This trend is interpreted to indicate that as the lower-frequency alternating pulse approaches the frequency of the higher-frequency pulse, stream segregation weakens and then disappears. Experimental group 6 (625 Hz alternating with 500-Hz pulse) has a generalization gradient that is most similar to that of control group 1, indicating that there was little or no essential difference between the perception of the alternating stimulus and that of the 625-Hz pulse stimulus repeated at 40 pps.

Figure 3C shows each experimental group mean gradient subtracted from the gradient of control group 2 (625-Hz pulse, 20 pps), providing an indication of the similarity between these respective gradients. Experimental group 6 (625 Hz alternating with 500-Hz pulses) shows a sharply negative difference that increases toward higher test repetition rates. This reflects the greater overall generalization of the 625–500-Hz group that was noted and potentially explained above. The large, absolute differences between these means and the control group 2 means (625-Hz pulse, 20 pps) at the higher test repetition rates are consistent with the interpretation that the perception of the 625–500-Hz alternating stimulus bore little resemblance to the 625-Hz, 20-pps stimulus, or that stream segregation did not occur in this case. The difference values for the 625–240-Hz gradient are all positive, and the function is nearly flat. The near independence of this difference gradient with respect to test pulse repetition rate is consistent with the interpretation that these two stimuli evoked a fundamentally similar perception. The positive values for the 240-Hz difference function

reflects the weak overall generalization of the 625–240-Hz alternating group in responding to the single, 625-Hz pulse trains. This effect was noted and potentially explained above. The other two difference functions are intermediate in shape and position, consistent with intermediate degrees of stream segregation.

## DISCUSSION

Goldfish were conditioned to alternating pulses (ABAB. . .) that differed in spectral envelope and then were tested for generalization to single pulse trains made up of the higher-frequency pulse (625 Hz) alone (AAA. . .). For the largest spectral difference between alternating pulse types, the generalization gradients were most similar to those produced by animals conditioned to the 625-Hz pulse alone, repeated at 20 pps, or one-half the overall pulse repetition rate used in conditioning. For the smallest spectral difference between alternating pulse types, the generalization gradient was most similar to those produced by animals conditioned to the 625-Hz pulse alone, repeated at 40 pps. Thus, the data are consistent with the notion that the widely frequency-separated alternating pulses evoked at least one segregated stream (corresponding to the 20-pps repetition rate), while the least frequency-separated pulses evoked only one stream (corresponding to the 40-pps repetition rate). This pattern of results is consistent with the hypothesis that goldfish are capable of a sort of auditory stream segregation similar to that observed in human listeners (e.g., Bregman and Campbell 1971). These results help confirm the conclusions of a previous experiment on auditory stream segregation in goldfish (Fay 1998a) in indicating that these perceptual behaviors are shared among humans, starlings (Hulse et al. 1997; MacDougall–Shackelton et al. 1998), and goldfish. The present results further support the hypothesis that the occurrence of stream segregation in goldfish depends on the spectral contrasts between sources: Stream segregation tends to disappear when the contrasts are small (e.g., 625- vs. 500-Hz spectral envelopes). Performance consistent with stream segregation declines monotonically as the spectral contrast between sources declines from the maximum to the minimum investigated here (see Fig. 3). The greatest decline occurs between the 625 vs. 240-Hz group, and the 625 vs. 305-Hz group.

The spectral envelopes of all stimuli of the present experiment overlap to some degree (see Fig. 1). It could be argued that generalization gradient differences simply reflect the degree of this overlap. In this case, for example, the growth in overall generalization that occurs as the lower-frequency pulse in alternating trains approaches the frequency of the 625-Hz pulse (see Fig. 3) could be ascribed to the increasing spectral

overlap of the alternating pulses. In other words, in spectral terms alone, it may have become harder to discriminate the alternating pulses from one another, and it came to be perceived that the conditioning and test stimuli generally have more elements in common, leading to more generalization. In this scenario, streams were not necessarily segregated; rather, the conditioning and test stimuli simply resembled each other more, overall, as the frequency of the lower-frequency pulse approached the 625-Hz pulse. While this hypothesis can account for the overall degree of generalization response strength in a simple way, it cannot explain why the fundamental shapes of the gradients transition from one consistent with a perceived repetition rate of 20 pps (the 625–240-Hz alternating stimulus—group 3) to one consistent with a repetition rate more like 40 pps (the 625–500-Hz alternating stimulus—group 6). This difference is consistent with the formation of at least one 20-pps stream in the former case, indicating segregation, and the formation of only one 40-pps stream in the latter case, indicating that segregation did not occur. Every group was tested with pulse trains that would be expected to form a single stream. Thus, if streams were segregated during conditioning, it would be expected that generalization to the single, simple stream would be greater than to the stimuli-evoking segregated streams. Under this hypothesis, the overall degree of response strength reflects the fundamental differences between conditioning and testing stimuli when the conditioning stimulus evokes a single stream at 40 pps and when the testing stimuli evoke one or more streams at 20 pps (less generalization—group 3) or another single stream equivalent to 40 pps (more generalization—group 6).

In general, the present and previous (Fay 1998a) demonstrations of auditory streams and their segregation in goldfish supports the speculation of MacDougall–Shackelton et al. (1998) that auditory scene analysis may be a common function of the sense of hearing among vertebrate animals. As Bregman (1990) emphasized, there really are distinct physical objects and events in an individual's environment, and that computations on the sensory code really ought to help create behavior appropriate to them. Scene analysis would seem to be a generally important process, serving such functions as predator and prey identification, interspecific communication and eavesdropping on other species' communication sounds, and in assessing the physical structure of the environment. In these cases, the source of a sound must be determined in order for appropriate behavior to occur, and something like stream segregation and scene analysis must play a role in this sound source determination. We believe that this most important function of hearing has played a major role in the

evolution of vertebrate auditory systems, and it provides the foundation required for effective, intraspecific acoustic communication and predator and prey recognition.

In the present experiment, stream segregation is investigated with respect to the temporal patterns and spectral envelopes of complex stimuli. Temporal pattern encoding and processing require crisp peripheral responses to waveform and envelope features. Spectral processing begins with peripheral frequency filtering. Both of these aspects of encoding and processing are the hallmarks of vertebrate auditory systems, and both have been investigated in goldfish. As in all vertebrates studied, the response of primary auditory afferents in goldfish precisely phase-lock to waveform fine structure up to at least 1 kHz (e.g., Fay 1978), and to subtle envelope events (e.g., Fay 1980; Fay and Passow 1982; Fay, et al. 1996). Psychophysical studies have shown acute temporal discriminations with errors as small as $50 \mu s$ for sinusoids (Fay 1978) and $200 \mu s$ for sinusoidal amplitude-modulated envelopes (Fay 1982). While psychophysical experiments have not been carried out using the pulses of the present experiment, the studies cited above suggest that the times between pulses are accurately represented in the peripheral neural code and with errors in interpulse time computation of $50–200 \mu s$ depending on pulse repetition rate. The differences in generalization gradient shape of the present experiment, as in (Fay 1995; 1998a), are probably determined by computations of interpulse time and to some degree reflect the temporal precision with which the pulses are encoded peripherally.

These periodic pulse trains also have a spectral structure made up of lines separated by the reciprocal of repetition rate. Although the goldfish auditory system has a set of peripherally tuned frequency filters (discussed below), they are rather wide (hundreds of Hz) and would not be expected to resolve the differences among the line spectra of the present periodically pulsed stimuli. Neurophysiological studies on peripheral spectral encoding in goldfish reveal at least two differently tuned channels of the saccular nerve, one tuned at about 200 Hz and the other(s) tuned at 600 Hz or above (Fay 1997). These frequency ranges were used in the present study to help specify the centers of spectral envelopes for the pulse types that would most likely result in stimuli affecting one or the other of the peripheral frequency channels somewhat independently (i.e., 240 Hz vs. 625 Hz). Since stream segregation is robust for sources with these spectral differences and declines for more narrow frequency spacing, it is probable that stream segregation is based, at least in part, on independent processing of inputs from these two frequency-selective channels. In this case, the ability to segregate streams in the present

experiment seems to be founded on independent temporal analyses within at least two frequency-selective peripheral channels. Thus, a peripheral analysis of the acoustic spectrum appears to be an important initial process in auditory analysis for fish, as it appears to be for all vertebrate species investigated (Fay 1992b).

These experiments on goldfish produced results similar to those expected for human listeners from experiments on perceptions of the same types of stimuli (Miller and Heise 1950; Bregman and Campbell 1971; van Noorden 1977). These previous studies determined that stream segregation becomes more robust as frequency contrasts grow from a mimimam value and tends to appear at the smallest frequency contrasts for more rapid tone pulse repetition rates. The repetition rates used in the present experiment are slightly higher than those used in human studies to define the existence regions for stream coherence (no segregation) and fission (always segregated) (van Noorden 1975). In addition, the tone durations used here are generally shorter than those used by other investigators of ABAB stream segregation (van Noorden 1977; Bregman and Campbell 1971; MacDougal–Shackelton et al. 1998). It is not possible to conclude that the perceptual experiences of human beings and goldfish are the same under these similar stimulus conditions. Nevertheless, the behavior of goldfish is consistent with a stream segregation hypothesis and has been shown to depend on the magnitude of spectral contrasts as it does for human listeners. All experiments on animal psychoacoustics face these validity questions when comparisons are made between humans and other species. Approaches to these problems depend on experimental replications using different methods and paradigms and on the accumulation of evidence for similarities and differences in the behaviors of different species, especially in the form of functional relationships. The present results thus may boost our confidence that there are many fundamental features of the sense of hearing that are shared among vertebrates, including humans.

In conclusion, fundamental features of the human sense of hearing, including aspects of auditory scene analysis, are shared among at least some other vertebrate animals. We suggest that auditory scene analysis is a very primitive function of hearing that has helped define "the state demanded by the organism" (Gans 1992) throughout the evolution of the vertebrate sense of hearing.

# REFERENCES

BEAUVOIS MW, MEDDIS R. Computer simulation of auditory stream segregation in alternating tone sequences. J. Acoust. Soc. Am. 99:2270–2280, 1996.

BREGMAN AS. Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press Cambridge, 1990.

BREGMAN AS, CAMPBELL J. Primary auditory stream segregation and perception of order in rapid sequences of tones. J. Exp. Psychol. 89:244–249, 1971.

FAY RR. Behavioral audiogram for the goldfish. J. Aud. Res. 9:112–121, 1969.

FAY RR. Auditory frequency generalization in the goldfish. J. Exp. Anal. Behav. 14:353–360, 1970.

FAY RR. Perception of amplitude modulated auditory signals by the goldfish. J. Acoust. Soc. Am. 52:660–666, 1972.

FAY RR. Phase-locking in goldfish saccular nerve fibres accounts for frequency discrimination capacities. Nature. 275:320–322, 1978.

FAY RR. Psychophysics and neurophysiology of temporal factors in hearing by the goldfish: Amplitude modulation detection. J. Neurophysiol. 44:312–332, 1980.

FAY RR. Neural mechanisms of an auditory temporal discrimination by the goldfish. J. Comp. Physiol. A. 147:201–216, 1982.

FAY RR. Hearing in Vertebrates: A Psychophysics Databook Hill-Fay Associates Winnetka, IL 1988.

FAY RR. Analytic listening by the goldfish. Hear. Res. 59:101–107, 1992a.

FAY RR. Structure and function in sound discrimination among vertebrates. WEBSTER D, FAY RR, POPPER AN. The Evolutionary Biology of Hearing. Springer-Verlag New York 1992b, 229–263.

FAY RR. Perception of temporal acoustic patterns by the goldfish (*Carassius auratus*). Hear. Res. 76:158–172, 1994a.

FAY RR. The sense of hearing in fishes: psychophysics and neurophysiology. Sens. Syst. 8:222–232, 1994b.

FAY RR. Perception of spectrally and temporally complex sounds by the goldfish (*Carassius auratus*). Hear. Res. 89:146–154, 1995. DOI:10.1016/0378-5955 (95) 00132-8.

FAY RR. Frequency selectivity of saccular afferents of the goldfish revealed by revcor analysis. LEWIS ER, LONG GR, LYON RF, NARINS PM, STEELE CR, HECHT-POINAR E. Diversity in Auditory Mechanics World Scientific Publishers Singapore 1997, 69–75.

FAY RR. Auditory stream segregation in goldfish (*Carassius a uratus*). Hear. Res. 120:69–76, 1998a. DOI:10.1016/S0378-5955 (98) 00058-6.

FAY RR Perception of two-tone complexes by goldfish (*Carassius a uratus*). Hear. Res. 120:17–24, 1998b. DOI:10.1016/S0378-5955 (98) 00048-3.

FAY RR, CHRONOPOULOS M, PATTERSON RD. The sound of a sinusoid: Perception and neural representations in the goldfish (*Carassius auratus*). Aud. Neurosci. 2:377–392, 1996.

FAY RR, PASSOW B. Temporal discrimination in the goldfish. J. Acoust. Soc. Am. 72:753–760, 1982.

FAY RR, REAM TJ. Acoustic response and tuning in saccular nerve fibers of the goldfish (*Carassius auratus*). J. Acoust. Soc. Am. 79:1883–1895, 1986.

GANS C. An overview of the evolutionary biology of hearing. WEBSTER DB, FAY RR, POPPER AN. The Evolutionary Biology of Hearing. Springer-Verlag New York 1992, 3–13.

HULSE SH, MACDOUGALL–SHACKELTON SA, WISNIEWSKI B. Auditory scene analysis by songbirds: Stream segregation of birdsong by European starlings (*Sturnus vulgaris*). J. Comp. Psychol. 111:3–13, 1997.

MACDOUGALL–SHACKELTON SA, HULSE SH, GENTNER TQ, WHITE W. Auditory scene analysis by European starlings (*Sturnus vulgaris*): Perceptual segregation of tone sequences. J. Acoust. Soc. Am. 103:3581–3587, 1998.

MCCABE SL, DENHAM MJ. A model of auditory streaming. J. Acoust. Soc. Am. 101:1611–1621, 1997.

MILLER GA, HEISE GA. The trill threshold. J. Acoust. Soc. Am. 22:637–638, 1950

TERHARDT E. Frequency analysis and periodicity detection in the sensations of roughness and periodicity pitch. PLOMP R, SMOORENBURG GF. Frequency Analysis and Periodicity Detection in Hearing. Leiden Sijthoff 1970, 278–290.

VAN NOORDEN LPAS. Temporal coherence in the perception of tone sequences. Ph.D. dissertation (unpublished). Institute for Perception Research, Eindhoven, The Netherlands. Cited in Beauvois MW, Meddis R (1996) Computer simulation of auditory stream segregation in alternating tone sequences. J. Acoust. Soc. Amr. 99:2270–2280, 1975.

VAN NOORDEN LPAS. Minimum differences in level and frequency for perceptual fission of tone sequences ABAB. J. Acoust. Soc. Amr. 61:1041–1045, 1977.

WISNIEWSKI AB, HULSE SH. Auditory scene analysis in European starlings (*Sturnus vulgaris*): Discrimination of starling song segments, their segregation from conspecific songs, and evidence for conspecific song categorization. J. Comp. Psychol. 111:337–350, 1997.