SSRs, then determinate colors are just determinate SSRs. In tying determinate colors to what normal perceivers can distinguish, B&H have, in their own words "failed to answer what we might call *Berkeley's Challenge,* namely, to explain why perceivers should be mentioned in the story about the nature of color, but not in the story about shape" (sect. 2.2, last para.).

Parallels between hearing and seeing support physicalism

Stephen Handel[a] and Molly L. Erickson[b]

[a]*Department of Psychology;* [b]*Department of Audiology and Speech Pathology, University of Tennessee, Knoxville, TN 37996.* **shandel@utk.edu merickso@utk.edu web.utk.edu/~aspweb/faculty/Erickson/default.html**

Abstract: There are 2,000 hair cells in the cochlea, but only three cones in the retina. This disparity can be understood in terms of the differences between the physical characteristics of the auditory signal (discrete excitations and resonances requiring many narrowly tuned receptors) and those of the visual signal (smooth daylight excitations and reflectances requiring only a few broadly tuned receptors). We argue that this match supports the physicalism of color and timbre.

The correspondences between the perceptual properties of hearing and seeing are not simply one to one, but one to many. Consider color: the intuitively obvious correspondence would be color to pitch. Each "pure" color and "pure" pitch can be associated with a single wavelength, and it seems natural to associate colors with pitches and vice versa. Moreover, although there are not complementary pitches or metamers, there are pitch intervals (octaves and fifths) that have unique perceptual relationships leading to the circle of fifths and spiral representations of pitch height (frequency) and pitch chroma (octaves) (see Shepard 1982).[1] However, we believe that a richer correspondence exists between visual color and auditory timbre.[2] Here color and timbre belong to objects. Color and timbre constancy allow perceivers to break the sensory world into coherent objects in spite of variations due to surface illumination or due to excitation frequency and intensity. Without source timbre, there would be no connections among sounds. We are using the term timbre in a nontraditional way. By the ANSI (American National Standards Institute) definition, timbre is that quality that distinguishes two sounds at the same pitch and loudness, and therefore, each sound-producing object produces a set of timbres across pitch and loudness. Yet, timbre must necessarily be a property of the source (e.g., a flute, a Barbra Streisand) that allows the listener to segment the varying acoustic signals into stable sources.

If we accept the match between color and timbre, then we can argue that there are fundamental parallels between the production of color and the color receptors in the retina, and the production of sound and the auditory receptors (hair cells) in the cochlea. Such a parallel does not prove that color is the spectra due to the surface reflectance, or that timbre is the spectra due to the sound body resonances. But the fact that the visual and auditory sensory systems are specifically "tuned" to the different type of sensory energy for each sense does buttress both contentions and weakens the argument that sensory qualities are arbitrary constructions.

Both color and timbre are conceptualized as source/filter models, although it is the fundamental differences between both the auditory and visual sources, and filters, that are crucial to our argument. What is common to both hearing and seeing is the independent "multiplication" of the source excitation energy by the filter response. At this point we can imagine a second source/filter process: the resulting frequency spectra becomes the source and the sensitivity curve for the receptors becomes the filter. The excitation of each receptor is based on the multiplication at each frequency of the filtered source excitation by the receptor sensitivity: presumably the firing rate is a function of that sum across frequency (see Fig. 1 in the target article).

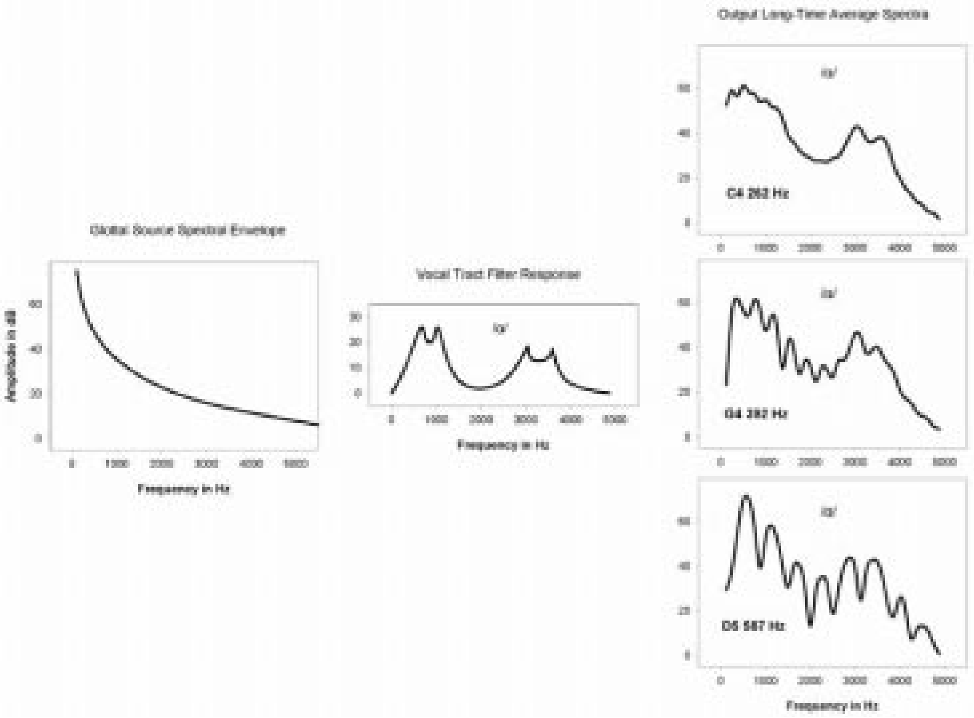

Figure 1 (Handel & Erickson). Representation of the source-filter model for the human voice. Output long-term average spectra are shown based on source frequencies of 262 Hz, 392 Hz, and 587 Hz.

Consider vision first. What we want to explain is why only a small number of cones are necessary. The source excitation will be due to direct sunlight, skylight, and reflected light from other objects, and the resulting excitation spectra of natural light at different times of day and locations is continuous and relatively smooth. Judd et al. found that the different excitations could be reproduced using different amounts of three independent functions: one function to represent the overall illumination level, one function to represent the blue-yellow contrast, and one function to represent the red-green contrast. The surface reflectance (the filter) is due to embedded particles that reflect the incident light. Somewhat surprisingly, the reflectance functions of most materials also are continuous and smooth, as illustrated in Figure 2 of the target article. Using diverse surfaces, most studies have found that the reflectance spectra can be reproduced with 3 to 7 independent functions (Wandell 1995) and that the first three functions usually represent (1) illumination, (2) red/green, and (3) blue/yellow contrasts. The fact that both the illumination and reflectance functions can be represented by a small number of independent functions suggests that only a small of number of receptors would be necessary to recover the illumination-independent color. However, even with three functions for both illumination and reflection there is not an explicit solution for trichromatic vision: there are six unknowns but only three data points from the cones. Maloney (1999) and Hurlbert (1998) present alternative simplifying assumptions that yield a solution for reflectance.

Now consider timbre. What we want to explain here is why there are roughly 2,000 sound receptors in the inner ear. The source excitation (e.g., bowing or plucking a violin, vocal fold vibration) occurs at discrete and typically harmonic frequencies, and the energy at each frequency depends on the precise ways the excitation is initiated. Bowing generates a different pattern of amplitudes than plucking, and the amplitudes of the higher harmonics are relatively greater at more intense excitation levels. The sound body resonances (the filters) also occur at discrete frequencies based on the shape and material of the sound body. In the case of the human voice, resonance peaks termed formants occur at frequencies determined by vocal tract shape and size, so the radiated sound usually contains multiple peaks at widely spread frequencies separated by regions of low amplitude (Fig. 1). What this means is that neither the source spectra nor the filter spectra can be modeled by a small number of independent linear functions, and timbre depends on the distribution of individual vibrations across frequency. To distinguish among different timbres (i.e., different sound objects) therefore requires many receptors, necessarily tuned to narrow frequency bands to pick up the resonance peaks; and that is what is found in the peripheral auditory system. The perceptual dimensions underlying similarity judgments between pairs of timbres are based on the amplitude pattern of the spectra. The dimensions include the spectral centroid (i.e., the weighted average of the frequencies), the number and frequency range of the harmonics, and the variance of the harmonics, particularly across the duration of the sound (Erickson, in press). All of these require a fine-grained analysis of the spectrum.

We believe this correspondence between the physical characteristics of light and sound and the characteristics of the visual and auditory sensory receptors support Byrne & Hilbert's (B&H's) contention that colors are physical properties, and support the analogous contention that timbres are physical properties.

NOTES

1. It is surprising that books rarely point out that sound waves are as "pitchless" as light rays are colorless. We suspect that writers are lulled by the correlation between frequency and pitch, which is not found for colors.

2. It is interesting that vocal pedagogues use the terms color and timbre interchangeably when referring to the quality of a voice (see Vennard 1967).

Byrne and Hilbert's chromatic ether

C. L. Hardin

Syracuse, NY 13210. **chardin1@twcny.rr.com**

Abstract: Because our only access to color qualities is through their appearance, Byrne & Hilbert's insistence on a strict distinction between apparent colors and real colors leaves them without a principled way of determining when, if ever, we see colors as they really are.

Hue differences are differences in quality. Spectral power differences are quantitative. This renders any putative identification of hues with spectral power distributions problematic. If the identification is to be made persuasively, it must be possible to show how hues – or hue magnitudes – can be mapped into spectral power distributions in a principled fashion. Byrne & Hilbert (B&H) propose to do this by relating hue magnitudes to relative cone response. For example, a light with a spectral power distribution that stimulates L-cones more than M-cones ("L-intensity") is to be denominated "reddish," whereas a light with a spectral power distribution that stimulates M-cones more than L-cones ("M-intensity") is to be deemed "greenish."

This talk of "L-intensity" or "M-intensity" sounds as if it were subject-independent, but it isn't. Not only do individuals differ in their opponent systems, the balances between opponent systems in a given individual are subject to shifts depending on luminance level, stimulus size and duration, and state of adaptation. If one could find a plausible specification of "L-intensity," "M-intensity," and "S-intensity" based on spectral power distributions alone, one could speak of the accuracy or inaccuracy of a person's visual estimates of hue magnitude, just as one speaks of the accuracy or inaccuracy of a person's estimate of length or weight. We can, indeed, measure the ability that people have to resolve wavelength differences precisely because we have an independent way to measure wavelengths. But without such an independent measure, it is simply nonsense to speak of the accuracy with which someone estimates hue magnitudes. All we can do is determine the extent to which people agree or differ in their hue magnitude estimates.

B&H attempt to blunt this sort of criticism by appealing to the well-worn distinction between something's being F and our ability to know or gain epistemic access to F. For example, in discussing simultaneous contrast, they distinguish between an object's *appearing* brown and its *being* brown. "If an object looks brown against a light background then it will look orange against a dark one" (target article, sect. 3.1.3, para. 1). However, "the fact that brown is only ever seen as a related color tells us nothing about the nature of brown. It merely illustrates the fact that color perception works better under some conditions than others" (sect. 3.1.3, para. 4).

So under what conditions does "color perception work better" (presumably, come closer to showing us the colors of objects "as they are")? Is there, for example, a background that is best suited for displaying the "true colors" of a set of Munsell chips? One would look in vain in the literature of color technology for an answer to such a question, not because it is hard to answer, or unanswerable, but because it is ill-conceived. As every practitioner knows, the choice of background is as much a function of one's purposes, as it is of the particular, empirically accessible, characteristics of the materials at hand.

Because they insist on a distinction between apparent colors and real colors, while acknowledging that access to color qualities can only be gained through color appearance, B&H are forced to a damning admission: "Thus we are prepared to countenance 'unknowable color facts' – that a certain chip is unique green, for instance. And so should any color realist who accepts some assumptions that are (we think) highly plausible" (target article, note 50).

There is at least a whiff of ether here, the electromagnetic ether whose undulations were supposed to be the mechanical basis of electromagnetic phenomena. The null result of the Michaelson-Morley experiment left one with two choices: Regard the earth's