Introduction

Vision and audition are two primary sensory modalities that inform us about the location and identification of objects in the environment. They are intimately connected to somatosensory systems, helping us to orient ourselves in the environment through balance and movement. Therefore vision and audition are essential to both motor learning and planning in the course of normal development.

In a companion monograph, I detail the parallels between auditory and visual processing. The concept of processing has been useful in differentiating components of sensory systems embodied in end organs, such as eyes and ears, from system centers in the brain and connections to the rest of the body. For the sake of convenience we sometimes refer to eyes and ears as the hardware of their respective sensory systems and their corresponding centers in the brain as the software of the sensory systems. As noted by Harris, the complexity of the neural networks involved in the human brain expose the computer model as an oversimplification.2

What makes us more intelligent than computers in processing auditory and visual information? We are not quicker or more precise, but we are superior at perceiving objects in natural scenes and noticing their relations. We have a greater ability to understand language and retrieve contextually appropriate information from memory. Understanding these distinctions is part of the appeal of the model of parallel distributed processing, helping to explain how we deal with concurrent streams of information through our auditory and visual systems.3

Another impetus for revisiting the significance of auditory and visual processing is that emergence, as a behavioral phenomenon, has re-surfaced in cognitive science.4 Behavioral optometrists are well-acquainted with the notion of vision as an emergent process, originally described by Skeffington as four interlocking subcomponent circles. The circles included centering, identification, vestibular/antigravity, and speech/audition, with the latter circle left relatively undifferentiated. Although left almost as a placeholder in his model of vision, Skeffington was on to something of deep significance when he included speech/audition in his concept of emergence. As noted by Kraskin, Skeffington’s terms are better conceived as descriptions of phenomenon rather than as physiologic activity. In other words, they are aspects of information processing as much as they are mechanisms, processes, systems, or subsystems.

Given that background, let’s review how select optometrists have dealt with auditory-visual integration and how current concepts in auditory and visual processing may lead to new clinical and institutional models.

A number of clinicians have had significant impact on drawing attention to the role of auditory processing in development and learning. I will review the historical context of the contributions of select clinicians and researchers, representative of key milestones in our field.

Jerome Rosner: Auditory and Visual Analysis

Dr. Jerome Rosner was in the private practice of optometry for twenty years before becoming a researcher at the University of Pittsburgh’s Learning Research and Development Center. In 1968 he co-founded the Pace School, a private school for children with dyslexia. His research led to the development and validation of an individualized perceptual skills curriculum, a landmark document published in 1972 and funded by the Ford Foundation and the United States Department of Health, Education and Welfare.

The pivotal tests that Rosner used in developing his perceptual skills curriculum were the Visual Analysis Test (VAT) and the Auditory Analysis Test (AAT). The VAT required reproducing dot-to-dot grid patterns, and the concept is still widely used in developmental optometry for vision therapy as G.O. board dot patterns. The AAT required phonemic awareness, typically parsing and manipulating the syllables in a word. The first item is “Say cowboy. Now say it again, but don’t say boy.” The last item, requiring more of a link between auditory discrimination and visualization is “Say smack. Now say it again, but don’t say /m/.”

Rosner re-worked the VAT to become the Test of Visual Analysis Skills (TVAS) and re-worked the AAT to become the Test of Auditory Analysis Skills (TAAS). This resulted in the perceptual skills curriculum evolving into the Preparation for Learning (PREP), and subsequently into two distinct, complementary programs: the Spatial Awareness Skills Program (SASP) and the Phonological Awareness Skills Program (PASP).

In the preface to the SASP, Rosner notes that although spatial awareness is a critical, developmentally derived precursor to elementary school achievement, it is not the only one. Phonological awareness skills are also important, and many children manifest deficits in both. That is why he wrote the companion PASP. Although the two programs differ in the skills they teach, they also have much in common. Rosner recognized that the SASP would be of greater interest to occupational therapists and developmental optometrists, with the PASP more appealing to speech pathologists and remedial reading specialists. Yet he firmly believed that all professionals should have an appreciation of how the two sets of skills interrelate.

An example of Rosner’s influence in phonemic awareness on optometric vision therapy is the Phonetic Focus procedure.
Figure 1. The Phonetic Focus procedure involving a reduced size near card with the beginning of the word, and a large chart for distance containing the end of the word.

adopted by Sarah Cobb (Figure 1). A nearpoint card contains single letters in columns on one side, and double letters in columns on the reverse side. A distance chart contains several letters in columns on both sides. The patient looks at the near card and blends the first letter with the corresponding suffix on the distance chart to form the word. In this manner the child is integrating a broad array of visual and auditory processing skills including near-far fixation, figure-ground, memory, chunking, word encoding, and sequencing.

Rosner’s footprints are evident in a manual published by the Department of Health, Education and Welfare, co-authored by Diana Phelps, a speech-language pathologist, and Rocky Kaplan, a developmental optometrist. There is significant emphasis on auditory and visual processing, and it served as a blueprint for the interdisciplinary clinic at the University of Houston. Parenthetically, the foreword to the manual was written by G.N. Getman, OD.

Vincett streamlined Rosner’s materials into an Optometric Perceptual Testing and Training Manual in the 1970s, with heavy emphasis on auditory and visual processing procedures for use at home as well as in the office. The visual processing material included G.O. Board and Parquetry Block patterns. The auditory processing material included auditory analysis (articulation sounds), auditory sequencing (blends), and auditory rhyming. Rosner published a guide for parents and teachers to use with children, with heavy emphasis on G.O. Board patterns for visual processing, and phonemic awareness followed by decoding skills for auditory processing. The material is outstanding, and an online preview is available at: (www.amazon.com/Helping-Children-Overcome-Learning-Difficulties/dp/143923180X#reader_143923180X).

Lane assembled an elaborated version of Rosner’s and Vincett’s approach in a manual with strong emphasis on therapy activities to develop motor, auditory, and visual processing. He includes an outstanding overview of the neurological complexity of reading, serving as the basis for these procedures.

Bowan took Rosner’s principles and merged them with Skeffington’s concepts to create a perceptual-motor model of language. He noted that phonological analytical skills are the entry point to reading and the ability to deal perceptually with the speech continuum is more important than phonemic data alone. Bowan related that one of Rosner’s reasons for his seminal research into perceptual remediation was his clinical frustration over children who were spatially competent yet learning disabled. His work led to the conclusion that the phoneme/grapheme relationship required a competency with analysis of space and time. This is the functional interface between visual/spatial performance and verbal/temporal performance, depicted by Bowan as a Venn Diagram with language as an emergent process of vision and audition (Figure 2).

Harry Wachs: Auditory and Visual Thinking

At about the same time that Rosner was developing his perceptual motor skills curriculum, another Pennsylvania optometrist, Dr. Harry Wachs, joined forces with fellow Piagetian scholar Hans Furth to publish Thinking Goes to School. Wachs considered visual-spatial thinking to be indicative of visual-cognitive intelligence. Furth and Wachs concentrated heavily on gross-motor integrative abilities together with auditory and visual cognition, though the auditory component is less apparent in Wachs’ subsequent work. The entire book is a treasure trove of therapy activities for processing and thinking in general movement, visual, auditory, and logic domains.

Erickson and Griffin point out that some of the basic activities in Thinking Goes to School rely heavily on visual-motor, auditory receptive, and speech-language expressive components of processing. The clap patterns, for example, involve auditory discrimination, auditory sequential memory, auditory visual integration, and most likely visualization.

There is evidence that Wachs purposely limited his approach regarding the speech-language aspects of auditory processing in pre-schoolers. In the manual of his pre-school cognitive tests, Wachs includes a non-standardized, supplemental section on auditory thinking. He qualifies this by writing that the evaluation of auditory thinking was not standardized because it was assumed that alterations in vocal presentation would vary amongst examiners. He goes on to offer the opinion that the growth of intelligence in some preschool children is actually impeded by an over-emphasis on language development.

John R. Griffin: Dysphonetic and Dyseidetic Dyslexia

Dr. John Griffin, based at the Southern California College of Optometry, spearheaded optometric involvement in the differential diagnosis of dyslexia. He proposed a neurological-behavioral model that clearly delineated the two principal forms of dyslexia: dysphonetic and dyseidetic.

In a subsequent paper jointly authored with Dr. Harold Walton, with whom he developed the Dyslexia Determination Test, Griffin made a crucial distinction about these two forms of dyslexia. They proposed that both dyseidesia, sometimes referred to as the visual subtype of dyslexia, and dysphoniesia, sometimes referred to as the auditory or phonological form of dyslexia, involve visual and auditory processing.

Due to failure to match visual and auditory gestalts of whole words, a dyseidetic individual may rely on subcortical vocalization to retrieve the meaning of a word. Conversely, a dysphonetic individual may rely on visual processing so that a whole-word visual gestalt can be processed quickly and matched with an auditory gestalt. Cortical vocalization, the inner voice that generates imagery, takes into account both phonological and visual processing. This involvement occurs in relative degrees, depending on the necessities of individuals with each subtype of dyslexia. A disruption at the level of
decoding prevents correct cortical vocalization, which results in poor word recognition, which then affects comprehension. Griffin’s influence in the field regarding linguistic components of auditory and visual processing is evident in two outstanding textbooks, one oriented toward generalized learning problems and the other toward reading in particular.

Harold Solan: Multimodal and Temporal Processing

Dr. Harold Solan was in the private practice of optometry prior to heading the Learning Disabilities Unit at the State University of New York (SUNY) College of Optometry, and becoming a prolific researcher at the College’s Schnurmacher Institute for Vision Research. Solan’s primary career interest was in reading, and he was heavily influenced by the work of Birch who hypothesized an orderly ontogenetic shift in sensory dominance from tactile to auditory to visual as a prerequisite to reading in the primary grades.

Birch was a physician in the Department of Pediatrics at the Albert Einstein College of Medicine and together with Belmont published an influential article on auditory-visual processing in the mid 1960s. Their findings led to the conclusion that deficits in auditory-visual integration contribute to incompetence in reading, and resulted in the Birch-Belmont Test of Auditory-Visual Integration (AVIT), a series of tapped sound patterns matched to their visual representation in dot patterns. It was his investigation of the AVIT as a temporal-spatial conversion as opposed to a task of intersensory integration that launched Solan into his research career in multimodal processing.

The differentiation of cognitive tasks into simultaneous vs. successive processing, a distinction in supramodal processing, also had significant influence on perceptual testing as well as therapeutic interventions. Linkages between cognitive processing and visual efficiency were noted by Groffman, who found a correlation of sequential function with saccadic processing and simultaneous processing with pursuits.

Solan was further influenced by the burgeoning research in temporal information processing that unified auditory, visual, somatosensory, and motor processes. These concepts were chronicled in a volume of the Annals of the New York Academy of Sciences dedicated entirely to the subject of temporal information processing in the central nervous system. Together with visual attention, coherent motion and other forms of temporal processing would come to predominate in a series of investigations on reading by Solan and his research group at SUNY. Solan established that children who lack competency in auditory-visual integration are at risk for experiencing difficulties in learning to read. Audition is subjected to language-based and temporal processing deficits; vision deficits are associated with spatial and temporal processing insufficiencies. Solan therefore concluded that reading disability is not necessarily a product of a deficit within either sensory domain. Rather, reading disability may be the result of a disorder in temporal processing common auditory and visual processing, commonly known as the magnocellular theory. Alternatively, the common linkage between visual and auditory processing deficits in developmental dyslexia may be impaired attention.

It is noteworthy that Solan and his colleagues pointed out the functional linkage of auditory and visual processing through the vestibular system as related to learning. They observed:

“During its course in the internal auditory canal that it shares with the cochlear nerve, the vestibular nerve also may affect hearing through efferent olivocochlear connections. Although either of these complementary dependent variables, vision and hearing, may function independently, together, they dominate our primary learning processes. The predominance of visual-vestibular control of balance gives way to a somatosensory-vestibular dependence by age three, but the transition to adult like balance responses is not complete for all sensory conditions even by age six. Since vestibular responses are associated with eye movements and hearing, they contribute to visual and auditory processing.”

Ken Gibson: Processing and Cognitive Enhancement

In 1985, Dr. Ken Gibson was an optometrist in Wisconsin who spearheaded an informal symposium involving experts in special education, clinical and cognitive psychology, occupational therapy, central auditory processing, visual processing, learning disabilities, and memory research.
from a number of universities and professional clinics. The composite model that they assembled for helping individuals with learning problems ultimately formed the core of Processing and Cognitive Enhancement (PACE). The processing components were auditory and visual processing, mediated by processing speed, attention, memory, logic, and reasoning (Figure 3).

Although Gibson has not as yet published his approach in peer-reviewed journals, it is a useful construct for auditory and visual processing as outlined in a self-published book about his proprietary program. In the area of visual processing, Gibson emphasizes visualization, imagery, manipulation, and visual thinking. In the area of auditory processing, phonemic awareness, sound blending, sequencing, auditory thinking, and verbalization are emphasized. His approach can be individualized and is well-structured and organized.

Precisely because Gibson’s approach is a pastiche of other approaches that have worked successfully for auditory and visual processing, elements of his program will be familiar to those acquainted with similar programs in the learning field, in particular Phono-Graphix and Lindamood-Bell. They all require a significant degree of rigor on the part of the instructor and self-discipline on the part of the student to attain mastery.

Consider this example of a procedure on spelling visualization, combining elements of auditory/linguistic and visual processing. Begin with words that are already familiar to the child, and that can be spelled individually. For instance, if the child’s name is Nancy, have her spell the sentence: My name is Nancy. The child spells each letter of the sentence individually in sequence without spaces: m-y-n-a-m-e-i-s-n-a-n-c-y. After that is completed the patient then alternates letters with the therapist. Once that can be accomplished the spaces are inserted between the words. However, this is done by silence visually representing the negative space, not by saying the word space. The spelling is done this way: m-y-[pause]-n-a-m-e-[pause]-i-s-[pause]-n-a-n-c-y. When the therapist alternates with the patient, processing speed is held at the threshold of the fastest pace that the child can maintain without getting lost.

**Burkhart Fischer: Auditory, Visual, and Optomotor Processing**

Professor Burkhart Fischer and his brain research group in Freiburg, Germany, have taken an approach toward learning in general and reading in particular that shares commonalities with Solan’s group. Fischer published a book in 2007 that summarized his approach to auditory, visual, and optomotor processing of children with learning problems. Fischer’s approach emphasizes non-linguistic learning skills in the auditory and visual domains.

Auditory discrimination, identified as low level or non-linguistic elements related to auditory processing, was subdivided into five basic components or features: volume, pitch, gap, time order, and side order (Table 1).

<table>
<thead>
<tr>
<th>Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Volume</strong></td>
<td>Measured with two white noise intervals 300 ms in duration. The interstimulus interval (ISI) was 150 ms. The reference signal was 55 decibels, and each trial started with a test intensity of 63 decibels. On each trial, the difference between target and reference stimulus was decreased by 10% of its previous value.</td>
</tr>
<tr>
<td><strong>Pitch</strong></td>
<td>Measured using a reference tone with 1000 Hz frequency, 300 ms in duration, and 65 decibel intensity. The test tone started with 1100 Hz (same duration and intensity as the reference tone). ISI was 150 ms.</td>
</tr>
<tr>
<td><strong>Gap detection</strong></td>
<td>Measured using 60 decibels, 300 ms white noise tones, one of which contained the gap. The two tones were identical in duration regardless of the gap. Gap duration started with 40 ms. ISI was 300 ms.</td>
</tr>
<tr>
<td><strong>Monaural time-order judgment</strong></td>
<td>Measured using a 1000 Hz tone and a 1120 Hz tone presented in random order.</td>
</tr>
</tbody>
</table>
Subjects were asked to indicate whether the higher tone or the lower tone was presented second. Both tones were 200 ms long and had an intensity of 63 decibels. The start value of the stimulus interval was 300 ms.

**Binaural side-order threshold** was measured through task clicks with 55 decibels delivered one to the right and one to the left ear in random order. Subjects were asked to indicate if they heard the second click from the left or the right. The start value of the stimulus interval was 300 ms.

Fischer and Hartnegg concluded that these auditory discrimination components involving temporal or magnocellular processing are significantly impaired in many dysexics. More importantly, these functions are amenable to training, and these training effects transfer to language-related phonological discrimination and spelling. 

The work of the Fischer lab in auditory processing serves as a complement to its results in low level or non-linguistic visual processing. This can be divided into optomotor processing principally involving fixation stability and saccades and visual perception principally involving subitizing. Fischer and Hartnegg concluded that deficits in antisaccade control contribute systematically to the problems of subjects with specific deficits in acquiring reading skills and that appropriate training can reduce the percentage of reading errors. I point this out because antisaccades, in which the subject is asked to look in the direction opposite to the stimulus location, are mediated by the pulvinar of the thalamus, and likely represents a common element in temporal order processing and visual attention.

Fischer and Hartnegg also found that dyslexics have a significantly higher percentage of binocular fixation instability, that there are commonalities with this population and those children who have ADHD and dyscalculia, and that these deficits can be trained. Unstable fixation reflects poor balance between parvocellular and magnocellular processing. Dyscalculia is a basic form of math learning disability, and Groffman demonstrated how the concept of subitizing can be assessed and trained through an interactive computer therapy program.

International awareness of Fischer’s work is promoted by a turn-key program he developed called BlickLabor in Germany and BlickMobil abroad. At the sixth International Congress on Behavioral Optometry in 2010, Peachey presented his experiences with Fix Train and Count Train, the visual processing components of Fischer’s program, and Fono Fix, the auditory processing component. Fix Train can be reinforced through procedures such as accommodative rock and pegboard rotator. Count Train or Subitizing can be reinforced through procedures such as geoboard and pegboard coding, and parquetry block patterns done as a tachistoscopic exposure activity. Fono Fix can be reinforced through rhythm and metronome procedures, Rosner phonemic procedures, and activities mirroring the Visual Aural Digit Span Test. The Visual Aural Digit Span Test consists of four subtests in which numeric sequences involving the digits 1 through 9 are used as stimuli. The child orally repeats or writes the sequences from memory. The first subtest requires oral repetition of orally presented digits. In the second subtest, the child must orally repeat digits that are presented visually. The third subtest requires the child to write digits that are orally presented. The fourth subtest involves writing digits that are visually presented.

**Keith Holland: Dual System Parallel Processing in Audition/Vision**

Keith Holland, an optometrist in the UK, has presented an approach toward parallel processing in audition and vision systems that shows the striking similarities between these two sensory systems. Virtually every function in the visual system has a parallel in the auditory system, from sensory receptors to neural connections, to what and where streams, and parvo/magno distinctions. There is transduction and transformation for sound in the cochlea analogous to processing that occurs in the retina. There is binaurality for localization of sound and space as there is binocularity for stereoscopic localization. There is top-down and bottom-up interaction of auditory processing in the inferior colliculus for sound as there is in the superior colliculus for vision. There is tonographic organization in the auditory cortex for acoustics analogous to topographic organization in the visual cortex.

From a developmental standpoint, auditory perception mirrors visual perception. The subcategories of figure-ground, discrimination, memory, sequencing, and closure are common to both domains, and attention can be a spotlight as well as a bottleneck. As we age, there is presbyopia as there is presbyopia, and if we suffer acquired brain injury there are auditory processing changes analogous to visual processing changes. Auditory neglect or semi-inattention to auditory space occurs analogous to hemi-neglect of visual space.

Holland’s unique contribution is his application of sound therapy to optometric practice. This is a fertile area for collaboration between optometry and audiology, as reviewed by the audiologist, Dorinne Davis. Davis points out the developmental sequelae of auditory deprivation, including sound mislocalization, hypersensitivity to sound, auditory processing timing lag, attention weakness, vestibular/balance disorders, and emotional instability that should resonate for developmental optometrists. She notes that vision therapy and sound-based therapy are complementary. A variety of sound based therapies are reviewed in detail in her book on the subject.

**Future Directions**

Much of the current involvement in using auditory processing concepts in vision therapy revolves around non-linguistic cues for cognitive loading or integrative purposes. Non-linguistic examples include the use of a metronome for pacing or timing, or the use of a game such as Simon for sequencing. Motor planning and timing accomplished through

---

**Table 1. Non-linguistic Low Level Components of Auditory Discrimination**

<table>
<thead>
<tr>
<th>Component</th>
<th>Functional Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>Identifying which sound is louder or quieter</td>
</tr>
<tr>
<td>Pitch</td>
<td>Recognizing which sound is higher or lower</td>
</tr>
<tr>
<td>Gap</td>
<td>Recognizing longer and shorter gaps between sounds</td>
</tr>
<tr>
<td>Time Order</td>
<td>Identifying which tone came first or second</td>
</tr>
<tr>
<td>Side Order</td>
<td>Identifying the order of two identical sounds heard right and left in random order (a form of binaural integration)</td>
</tr>
</tbody>
</table>
auditory-visual integrative activities such as Interactive Metronome can be valuable in developing attention.52 Cognitive loading can be done by having a patient perform a mental operation while engaged in a therapy procedure, such as adding numbers while performing spatial localization tasks.

While useful, these non-linguistic procedures may still leave learning and reading disabled children unable to transfer acquired skills to classroom and homework performance. This was Rosner’s original motivation in emphasizing the linguistic and lexical features of auditory processing and therapy.53 The ultimate transfer occurs when these approaches are combined into “smart schools” as reviewed by Lemer.54 A model that integrates auditory and visual processing into a school curriculum was pioneered by Ingersoll and described as integrated visual learning.55

For the foreseeable future, select optometrists will continue to push the envelope of programs that synthesize the linguistic and non-linguistic components of auditory processing together with visual processing. Although the interface between auditory and visual processing extends to many areas of practice, from early intervention services to acquired brain injuries, the most fertile area of application remains learning disabilities.

Jill Stowell is a special education teacher who has written a book that is a contemporary roadmap for parents seeking guidance about learning disabilities.56 In her acknowledgements she singles out Dr. Samuel Berne who helped her understand and organize neurodevelopmental, motor, and visual skills training. She also acknowledges the influence of optometrists, William Bescoby, Ken Gibson, Doug Stephey, and Al Sutton. Her chapters include core learning skills centering on visual and auditory processing. This type of effort is welcome in promoting awareness of the ongoing trans-disciplinary contributions of optometrists.

Salus University is an institution of higher learning in Pennsylvania that houses both a College of Optometry and a College of Audiology.57 Its program description makes it apparent that it is at the nexus of biomedical science and holistic care regarding visual-vestibular interaction in posturography and dizziness and balance disorders. I can visualize the University integrating its focus on vision and audition as related to learning, in the state where Rosner and Wachs began their journey.

References
tal/search/detailmini.jsp? _rft_t=true& _RERIEExTextSearch_Step=7 SearchValue 0-ED063098&ERICExTextSearch_Type=0-no&accno=ED063098 Last Accessed March 8, 2012.
36. Lallier M, Valdois S. Sequential Versus Simultaneous Processing Defici
cits in Developmental Dyslexia, Dyslexia - A Comprehensive and Inter


Corresponding author:
Leonard J. Press, OD
Family Eyecare Associates
17-10 Fair Lawn Ave.
Fair Lawn, New Jersey 07410
pressvision@aol.com
Date submitted for publication: 14 May 2012.
Date accepted for publication: 20 June 2012.