A Multisensory System for the
Development of Sound Awareness
and Speech Production

Kimberly V. Fischer and Barbara J. Parker
Department of Communication Sciences and Disorders
University of Oklahoma Health Sciences Center

The Multisensory Sound Lab (Oval Windows Audio) is educational technology developed for use by groups of children who are deaf or hearing impaired. It provides simultaneous visual, tactile, and auditory information in response to acoustic inputs. A sound-sensitive floor that is large enough to seat 15 children provides the vibrotactile information. Frequency information may be transposed downward as much as 2 octaves in order to enhance the vibrotactile and auditory experience. Visual feedback includes colorful light spectra, kaleidoscope patterns, and a frequency sensitive light tower. The amount and type of multisensory reinforcements can be adjusted by the user. Sound lab applications might include auditory detection, sustained voice production, or speech sound production activities. This paper describes applications, advantages, and limitations of the Multisensory Sound Lab for use in a clinical setting.

Traditionally, auditory skills have been described as a multi-stage process (Erber, 1982). First, detection or attention to sound is required. That is, only recognition of the presence or absence of sound is necessary. No categorization or meaning is linked to the sound. The second stage is discrimination between sounds. Differences among speech sounds may be very small, yet carry a great deal of information, such as the simple addition of /t/ to the end of a word. The third step, identification, requires the listener to understand that a sound is “different” and also to attach some memory to the sound. Comprehension, the final stage,
is the most difficult. The linguistic and non-linguistic coding that is embedded in the acoustic signals must be understood (Condon & Sandt, 1974; Engle, 1973; Hirsh, 1980).

More recently, Aslin and Smith (1988) proposed a developmental model in which sensory skills emerge sequentially in three levels: the sensory/primitive level, perceptual representation level, and the cognitive/linguistic level. Moeller and Carney (1993) further describe how neural auditory/perceptual skills may develop within this model. First, the sensory/primitive level develops as the child's sensory systems change incoming physical signals into a stimulation pattern. In the auditory system, the sensory/primitive level includes sound detection. Next, the child uses the stimulation pattern to form a complex neural code or perceptual representation. This level is thought to address pre-linguistic aspects of speech perception that have been observed in infants (Eilers, Wijnen, & Moore, 1977; Kuhl, 1983; Spring & Dale, 1977). The third, and highest level of representation occurs at the cognitive/linguistic level. Auditory comprehension of words and longer utterances appears at this level. Aslin and Smith's developmental model also has been extended to development of speech production in children with hearing impairments and children receiving cochlear implants (Carney et al., 1993; Moeller & Carney, 1993). Helping children who are hearing impaired function successfully from sensory/primitive throughout linguistic stages is important for the development of speech perception and production.

A number of aural rehabilitation programs can be used to address auditory skills described in traditional and/or developmental approaches (Erber, 1982; Shea, 1992; Stout & Wissler, 1992). Where appropriate, such programs make use of combined tactile, visual, and auditory information to help people learn to detect, discriminate, identify, and comprehend speech. Further, recent reports suggest that the use of multi-modality stimulation (visual, auditory, and tactile) may provide better results than the observed with only one or two sensory modalities (Carney & Beathler, 1986; Carney, Dunkel, & Beathler, 1981; Pratt, Heinitzman, & Deming, 1993; Roberts, Osberger, Miyamoto, Renshaw, & Carney, 1988). Technologies also have been developed to provide immediate multisensory feedback for speech production. When used to assist individuals with hearing impairment, devices are intended to provide the speaker with additional visual or tactile information related to their own motor or acoustic speech events (Fletcher, 1989). In some cases these include computer-based systems (Osberger, Moeller Kroeste, & Lipprann, 1981; Pratt et al., 1993). Most of the currently available systems provide simultaneous feedback in one or two sensory modalities.

The purpose of this report is to describe trial applications of the Multisensory Sound Lab (Oval Window Audio). The system, originally designed for use by
groups of children who are deaf, provides simultaneous sensory stimulation through visual, tactile, and auditory modes. Educational applications representing a variety of primitive and linguistic levels of intervention are provided. The examples include both single-user as well as group intervention methods.

The Technology

The Multisensory Sound Lab is an audio system that amplifies sound through loudspeakers and simultaneously transforms sound into floor vibrations that can be felt through the body and seen through a variety of visual displays. Typical components of the system are shown in Figure 1. The components can include a vibrotactile floor, spectrum analyzer (Visualizer), frequency sensitive light tower (Luna Light), laser light display, tone generator, oscilloscope, audio speakers, and induction loop assistive listening systems for the hearing impaired. Multiple sound sources such as microphones, audio cassette, CD, or VCR signals provide input to the system. Electronic musical instruments and games are especially well suited to the system. For example, battery-powered sound games (e.g., "Simon") synthesizers (e.g., electronic keyboards and drum sets), or electronic stethoscopes, and an electronic bass guitar are often used with the Sound Lab.

The sound floor (see Figure 1(a)) is comprised of nine interlocking modular panels, each of which covers a 16 ft² (1.5 m²) area. (Figure 1 includes only four panels in order to conserve space.) Once assembled, the floor covers an area up to 144 ft² (13.7 m²). The vibrotactile floor is driven by two loudspeakers that are coupled to the corners of the floor area (see Figure 1(b)). People sitting or standing on the floor will perceive sound as vibrations. High frequency sounds that are outside the floor's 20 Hz - 500 Hz range may be transmitted downward as much as 2 octaves in order to enhance the vibrotactile experience. Two loudspeakers (in addition to those that drive the floor) are provided so that users have greater control over audio listening levels (see Figure 1(c)).

The Visualizer (see Figure 1(d)) is a 6 octave spectrum analyzer (8 kHz full scale) that displays harmonic content as colorful vertical bars on a television monitor. The type of display is controlled from a keyboard (see Figure 1(e)). Special information is displayed in real time. The system does not provide for data storage, but instantaneous spectral peaks can be captured and held temporarily on the monitor screen. Harmonic content also may be displayed in color on a video kineidoscope patterns (see Figure 2), a feature useful for small children.

The patterns are sensitive to both the frequency and periodicity of signals, similar to tachistoscopic figures created when sine wave voltages of various amplitude and phase relationships are applied to the X and Y modals of a cathode-ray oscilloscope. Quasi-periodic signals, such as vowels, create an open circle or square pattern (see Figure 2a). Quasi-random signals, such as fricatives, result in filled or

can be obtained from Norman Loderman, Director, Oval Window Audi, 33 Wildflower Ct, Nederland, CO 80466; Voice, Fax, or TDD (303) 447-1007.
Figure 1. Example components of the Multisensory Sound Lab include: (a) sound floor, (b) front loudspeakers, (c) audio loudspeakers, (d) Visualizer, (e) keyboard, (f) Luna Sound Light, (g) light controls, (h) caisson, (i) cassette tape player, and (j) control box. Sound toys and a laser light display (not shown) also are available with the system.

Figure 2. The kaleidoscope display produced by the Visualizer for a vowel and fricative, respectively.
snowflake patterns (see Figure 2b). Larger figures correspond to lower frequencies; smaller figures correspond to higher frequencies.

Another visual display is the Luma Sound Light (see Figure 1f). It is a 7 ft (2.1 m)-tall column made of translucent plastic that contains three banks of color lights. The color lights respond to different frequency bands and sound intensities. The sensitivity of each color bank can be adjusted independently so that the user can modify the relative amounts of visual feedback for different frequency bands (light controls: see Figure 1g). With the ambient lighting dimmed, the Luma Sound Light provides a larger-than-life, colorful, and attention-gaining reinforcement. A dual storage oscilloscope (see Figure 1b) also is included to allow for waveform display and calibration.

The Sound Lab includes separate amplifiers for up to five simultaneous external inputs. In our system, one of these amplifiers is dedicated to a cassette tape player (see Figure 1f). Additional amplifiers provide for attenuation of floor vibration and light sensitivity separate from the external input gains. This allows the attenuation or amplification of visual reinforcement to be independent from the acoustic signal conditioning. For example, a clinician can gradually attenuate the amount of visual and vibrotactile reinforcement while maintaining the necessary sound audibility.

Because the Lab sustains use by many student clinicians, outside personnel, and groups of clients, it is, by necessity, sturdy and easy to use. The signal input, control electronics, amplifiers, signal generator, and frequency counter are housed within a single box that is covered with Velcro (see Figure 1i). Small peripheral toys and a sound level meter adhere to the box with Velcro for ready availability. For ease of use, red “don’t touch” stickers are used to mark the default settings for most applications. Any button or knob without a sticker is free for manipulation by the user. An informal survey of our student clinicians and clinical faculty members revealed that about 45 min were needed to become proficient in operating the equipment and accessory devices. In 2 years of operation we have experienced only one minor equipment failure.

**EXAMPLE: APPLICATIONS**

**Application 1: Sound Detection**

A non-verbal 2-year-old female with bilateral moderate to severe hearing loss was introduced to the Sound Lab. She was accompanied by her mother and clinician. A primary concern was soteresponsiveness to sound. The child wore binocular hearing aids and was seated on the sound floor with the assistance of her mother. Environmental lights were dimmed to minimize distraction and maximize the visual impact of the Luma Sound Light. Sound stimuli included the mother’s voice, an electronic bass drumbeat, and a battery-powered toy cow that “mooed.” The mother’s voice and cow’s moo were directed to the system.
with two microphones, and transposed downward to maximize the vibrotactile sensation of the floor. Each sound was presented at suprathreshold level and without maximum feedback from the floor and Luma-Light. The clinician monitored the child's responses to the paired sound and multisensory reinforcement. Head turns, pointing, and foot touching were interpreted as positive responses. After consistent response to the paired stimuli, the feedback from the floor and Luma-Light gradually was reduced.

When sound was first presented, the child immediately lay upon the floor, palms down, in front of the Luma-Light. This behavior presumably increased the child's sensation of floor vibration, although there are relatively few systematic studies of vibrotactile sensitivity on areas such as the highs, toes, and chest. Within a 20-min period, however, the child responded consistently to the cow's moo and her mother's voice without visual or vibrotactile feedback. With no simultaneous multisensory stimulation, occasionally the child evidenced frustration and pointed at the still Luma-Light when sound was presented. In this case, the clinician simply flashed the Luma-Light as reinforcement for correct detection of sound.

Application 2: Sound Comprehension

Five kindergartners and first graders who are moderately to severely hearing impaired were referred for trial classroom activities in the Sound Lab. One behavioral objective for these children was naming (oral or manual) of soft and loud sounds. The children were accompanied by two teacher assistants and a speech-language pathologist. All were seated in a circle in front of the Luma-Light and television monitor. All children wore binaural hearing aids and were reported to benefit from their use. The teacher assistants indicated the none of the children consistently identified both loud and soft sounds in the regular classroom.

The clinician alternately placed hard or soft objects in a can and shook the can in front of a microphone that provided input to the Sound Lab. Objects and amplifier gains were selected such that hard objects resulted in visual, auditory, and vibrotactile feedback. Soft objects, however, resulted in only auditory feedback. The auditory signal of soft sounds also was somewhat attenuated, though audible, for the children. Objects included were marbles, pennies, hard plastic toy figures, wet sponge pieces, crumpled paper balls, and foam squares. The can was passed among the children who each shook the can in turn and would sign and/or vocalize the concepts "loud" or "soft."

After children were familiar with the sound system and identification task, the clinician would alternately place hard or soft objects in the can. She then gave an individual student the opportunity to shake the can and identify the sound as "loud" or "soft." Other children would monitor the accuracy of their prior's answer. Near the end of the activity, the clinicians put cotton balls into the can.
this case, can shaking resulted in no audible sound, and no visual or vibrotactile floor feedback. When faced with this dilemma, children would shake the can vigorously while scrutinizing the possible sources for feedback (floor, light, speakers). After several attempts, they would shuffle their shoulders or demand to see if there was anything in the can. After 15 min with the activity, one child discovered that she could make soft or loud sounds with the same object by changing the vigor with which she shook the can. Other children initiated this experimenting behavior. In the regular classroom, the teacher and children subsequently completed a similar activity using auditory trainers without the additional vibrotactile or visual feedback.

Application 3: Gross Vocalization

A 5-year-old male who is severely hearing impaired was referred for trial activities in the Sound Lab. A primary concern was his inability to maintain voicing. Clinicians and teachers had encouraged the child to place his hand over the larynx in order to tactually sense voice vibrations, a process that yielded occasionally successful but inconsistent results. The child was accompanied to the Sound Lab by four peers who are hearing impaired and a classroom teacher. All children wore personal FM systems throughout the activity. Elbow-held microphones were used for sound input. Frequency information above 800 Hz was attenuated so that blowing, loud breathing, glottal clicks, voiceless fricatives, and other non-target noises would not register visual or vibrotactile feedback. The male client stood in front of the television monitor. Spectral information was displayed in a kaleidoscope mode so that the modeled vocalizations made open squares or circles on the screen. Two microphones were passed among the other children who each vocalized in turn and lit the screen. Children were encouraged to light the screen as long as possible, thereby prolonging vocalization.

Next a microphone was provided to the client. The clinician elicited a brief instant of voicing from the child using modelling and placement of the child’s hand over his larynx. After six attempts with simultaneous visual feedback, the child was able to maintain voicing for several seconds. The vocalization was quickly generalized from sustained vowel-like productions to alternating open-closed syllable-like strings. After having achieved success, the child was unwilling to yield the microphone so that other children could take turns.

Application 4: Articulation

Six third and fourth graders with moderate to severe hearing loss were accompanied to the Sound Lab by their teacher and speech-language pathologist. All children wore personal FM systems for the group activity. One concern for these children was an inability to contrast /s/ and /f/ production. For these children, the Lumas Light was calibrated so that the frequency information of /s/ and /f/ (produced in isolation and syllables) triggered differently colored lights. Spectral
feedback from the Visualizer also was used to differentiate the frequency components of the two phonemes. Two microphones provided system input. The audio input was transposed downward approximately 2 octaves, enabling the /f/ production to be sensed minimally through the vibrating floor. Correct /h/ production did not result in floor vibration that could be felt.

First, the clinician modeled correct production of /h/ and /f/ in isolation. Children were shown the vibratotactile and visual differences between the two phonemes. After children could identify the sound from its multivibrator feedback, they were encouraged to produce one of the phonemes following the clinician's model. Cued speech was used to maximize the effectiveness of the model. Children took turns, with one of two microphones. All children monitored the production to determine whether it was more like /h/ or /f/.

The children quickly recognized that they must change their production in order to distinguish between the two phonemes and identify the correct phoneme. Frequently, however, their efforts resulted in further distortions rather than the original attempt. For this reason, it was necessary for the clinician to model and cue each production. Because fricative production is relatively low intensity, it was necessary for the clinician to maintain close mouth-to-microphone distance for the children. Cloth chair covers helped to stabilize the active child and improved the success of the activity. Small clip-on microphones (also available with the Multivibrator Sound Lab) could be used with a head mount for activities requiring very close, or highly stable mouth-to-microphone distance.

Application 5: Music Enrichment

A music teacher was concerned that her fourth through sixth grade students who are hearing impaired could not match the meter or tempo of music with body movements or clapping. The teacher had been working on this skill by having students feel the piano sound board or use a Metronome. She reported that these activities had limited success. All children were severely or profoundly hearing impaired and wore bilateral hearing aids.

We asked the teacher to select an audio tape that she used in the classroom. In addition, low frequency dominant stimuli (a tape of tap music, and tapes of common folk songs played in callo) were provided. An electronic bass drum and non-tonal rhythm were produced by a synthesizer. The tempo could be adjusted continuously. A heart beat also was transduced using an electronic stethoscope and卖给 student participant. The cassette tapes, electronic drum beat, and heartbeat provided the input to the sound system. Stimsi with both predominant beat and low frequency components were considered the easiest to follow rhythmically. The stimuli were presented in order from least to most difficult. The first stimulus was the electronic drum, followed by the heartbeat, tap music, folk songs, and the familiar classroom tape.
Seven students sat on the floor and attempted to match the beat with a foot up. The teacher assisted some students by patting the basic meter on their shoulders. When each student successfully matched the tempo, a new stimulus was selected. Soon students were familiar with the stimuli as well as concepts of “basic beat” and “tempo.” At this point, each student was given a musical instrument (drum, tambourine, electric bass guitar, xylophone, or cymbal). The musical sounds were simultaneously directed to the system using external toy inputs and microphones. Gains were adjusted such that high frequency sounds by the Luna Light, but low frequency sounds created floor vibration. The students instrumentalists were arranged with high frequency sounds near the Luna Light. A conductor was selected from the group to direct the musicians. The conductor could direct their own unique musical piece, or select from the audio-taped stimuli.

Both students and teacher were enthusiastic about the activity. Students were especially interested in the kick drum and instruments. However, by the end of the 45-min session, at least 3 of the 7 students remained unsuccessful in their ability to match the tempo even with the additional visual and vibrotactile information. Often children become more interested in their own music making, than in matching the tempo of others. Still, this activity demonstrates the degree to which a creative, musical, auditory experience can be enriched with visual and vibrotactile information.

Summary

Numerous hardware and software products are now available to provide visual and auditory interactive feedback for auditory habilitation or speech-language therapy. Frequently these products are limited in application to a single user for individual therapy or self-instruction. Service delivery, however, often involves small group or classroom instruction. In this case, technology that is designed for group use can be cost effective. The Multisensory Sound Lab seems to meet this need. Although there are a wide variety of products that can provide reinforcement or interactive feedback in one or two sensory modalities (e.g., IBM Speechviewer, Kay Visips), the Multisensory Sound Lab provides simultaneous visual, auditory, and tactile feedback. This redundancy potentially could help children with a variety of perceptual problems to utilize their sensory strengths.

Clinical intervention that targets gross sound discrimination and identification, or production of speech targets differing in gross duration, intensity, or frequency information is well suited for the Multisensory Sound Lab formats. One limitation of the Sound Lab is the relatively poor frequency resolution of the Visualizer. The 6 octave band spectrum, for example, in 1,000 bands clearly demonstrate differences between vowels. The low cost and simple operation of this component, however, make this limitation less serious.

The Multisensory Sound Lab is not immune to problems that plague clinical application of special analyses in general. Soft, high-pitch, or hypernasal ph-
nation can produce inadequate or inaccurate visual feedback for the client. The /s/ and /f/ activity described here was limited to gross pinnate approximation in an isolated context because spectral content of fricatives is highly speaker and context dependent (Harris, 1994; Hughes & Hall, 1996). Furthermore, clinicians must evaluate the saliency of visual information provided in the instantaneous spectral context of fricatives. For example, clinicians should expect the noise floor limits of /s/ and /f/ to overlap (Stevens, 1969). Also, classification of /s/ distortions from spectral features may not be possible in many cases (Domlof, Wilcos, & Stephens, 1980). This complexity can decrease the ease with which children extract important information from the visual display. Fine discrimination and production activities can be addressed using other techniques.

A semi-permanent space should be allocated for the laboratory because the floor is not easily moved after assembly. The modular, 1.6 m² (1.8 ft²) interlocking floor panels allow flexibility in classroom arrangement. A niche panel, 13.4 ft² (12.37 m²) floor space is suitable as a room comfortably. Client groups larger than 15 are seldom seen in our clinic, so the sound floor dimension has not been a limiting factor. Because relatively few producers make extensive use of tactile information for learning (with the exceptions of vibratory and hearing aids), the vibrating floor seems to be the most innovative component.

The cost of a system ranges from approximately $3,000.00 to $15,000.00 and depends upon the variety of components selected by the user. A system similar to the one described here currently lists for $13,000.00. Individual components may be purchased separately, so the user can add to the laboratory as finances allow. The sound floor is a relatively more expensive component and costs approximately $395.00 per panel.

Our experience with the Multisensory Scint Lab has been positive. Faculty and students from other universities, private speech-language pathologists and audiologists, and public school personnel often use the lab for a variety of activities. These activities have involved teaching acoustics, speech therapy, vocal hygiene, hearing conservation, music therapy, and deaf education. Multisensory activities may enhance even the most bored or therapy-weary client. Given access to such versatile technology, the creative clinician and classroom teacher might develop a wide variety of activities that will motivate a variety of clients (Binter, Gilmore, & Fehry, 1994). Although initially promising, it will be necessary, through systematic study, to demonstrate the efficacy of this form for specific applications before it can be recommended as a therapeutic tool.

Acknowledgements

We express our thanks to those anonymous reviewers and Patrici Klicos, Ma., for their thoughtful comments. Also, we gratefully acknowledge the technical assistance of Norman Leekam and Buffalo Hearing. The site utilization equipment was kindly provided by Oral Windows Aid.
REFERENCES


