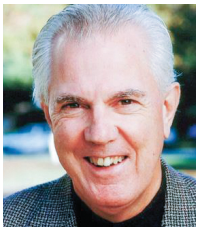
Serendipitous Observations in the Study of Auditory Perception

William A. Yost



Serendipity, what is it? To me, serendipity is the luck of finding or creating interesting or valuable things somewhat by chance. I believe I have experienced such serendipitous luck, but I do not think serendipity for me, as a scientist, occurred randomly. "Creating interesting or valuable things" does not occur in a vacuum. The "creation" is likely due to a whole host of preceding events that the scientist has been studying and thinking about. Of relevance for this essay is the idea of a zeitgeist. Zeitgeist ("spirit of the age") is a German term that to me describes an invisible idea dominating the characteristics of a given epoch in history. The "characteristics of a given epoch in history" can set the stage for a serendipitous creation.

Several of my serendipitous moments occurred when I realized that a current way of studying auditory perception probably needed to change and there was a "zeitgeist" that I experienced suggesting what the change might be. One of those moments of serendipity occurred when I realized that a great deal of the history of the study of auditory perception did not deal with what people actually say they "hear." Instead, a prevailing view of sensory perception after the nineteenth century was that the mind (brain) gains knowledge of the world based on the brain's ability to determine objects in the world via the senses (e.g., see Boring, 1942). It was not long before studies of auditory perception investigated the perception of the acoustic properties of sound (e.g., frequency, intensity, spectrum, duration, location) and the attributes associated with those acoustic properties (e.g., pitch, loudness, timbre, perceived duration, perceived location).

In 1989, I intended to write a textbook on sensory processing, explaining this history based on what I had learned as the Program Director for the Sensory Physiology and Perception Program of the National Science Foundation (NSF) in the early 1980s. This NSF program was responsible for funding any meritorious proposal-seeking support to investigate the biology and/or perception associated with any sensory system, and I learned a lot about sensory systems in this challenging job.

There were, of course, already such sensory-processing textbooks, some of which were very good. However, they were usually not "built" around a theme, and students often complained along the lines of, "Well it must be Tuesday, as I see the next topic is smell." I thought I had a theme based on the history I mentioned above that the brain uses sensory systems to process objects in the world and their attributes. I outlined chapters about vision based on light produced and reflected by objects and the perceptual attributes of this light.

However, when I started to do the same thing for hearing, I had my serendipitous moment in which I realized that almost nothing was known about objects that produce audible sound (the human voice and musical instruments were exceptions) or what perceived sound indicates about the objects that produce sound. I realized that most, if not all, people could, if asked, indicate if the pitch of a sound is low or high or if a sound is loud or soft. However, that is not what is usually voluntarily reported when one indicates what one hears. Instead, listeners often report

More from this author on Across Acoustics

LISTEN

"What Is Silence?"

bit.ly/AA-silence

https://doi.org/10.1121/AT.2024.20.4.81

the object (*source*) that produced the sound, such as the *car* screeched, the *baby* cried, the *ball* bounced, or the *leaves* on a tree rattled. Thus, auditory perception has a lot to do with the sources of sounds, and little seemed to be known about the perception of sound sources per se. In addition, how do we possibly sort out what the objects are when more than one object produces sound at about the same time?

Based on this bit of serendipity, I immediately stopped work on a sensory sciences' textbook and wrote an article, *Auditory Image Perception and Analysis* (Yost, 1991), explaining my serendipitous conclusion that "hearing" was about the objects that produced sound, not just about the sound itself. I am deeply indebted to Dave Green, as he gave me good ideas and the confidence to try to publish a nonresearch-based opinion piece.

Of course, I was not the only one making this observation at this time, with the most compelling of those being *Auditory Scene Analysis* by Al Bregman (1990), a book that clearly captured the imagination of sensory and perceptual scientists. There was obviously a zeitgeist in the late 1980s early 1990s because none of us who wrote about our serendipity at this time were aware that others were also having the same ideas. The challenges of understanding the perception of sound sources became the topic of several more of my publications, the organizing theme of all my research after 1990, and the generator of ideas for most of my experiments.

As I indicated previously, my efforts to understand sound-source perception had a profound influence on what I thought was important to know about auditory perception. This was during a period that followed major new discoveries of how the inner ear and auditory nerve process sound (e.g., Brownell, 2017). Consequentially, studies of auditory perception seemed to be related mainly to the peripheral processing of sound.

When I moved to Arizona State University (ASU), Tempe, I had the opportunity to work with Michael Dorman and those in his laboratory on a series of papers investigating cochlear implant (CI) patients' performance in different spatial-hearing tasks (see Pastore et al, 2024, for the most recent paper). The success of the CI demonstrated to me that the auditory periphery (an "ear") is not required to perceive sound but a *brain* is.

Indeed, a type of serendipity occurred when I realized sound-source perception was about *what* and *where* sound sources are, and the *brain*, not the auditory periphery, determined *what* and *where* sound sources are. As I developed these ideas about sound-source perception, I also realized that to perceive sound sources, it was not necessary to be able to identify (label) *what* a sound source was. When Bob Lutfi wrote a chapter (Lutfi, 2007) for the book on sound-source perception that I coedited (Yost et al., 2007), we argued about the extent to which sound-source identification was important. Bob is an astute scholar of auditory perception, so his ideas about sound-source identification stayed with me.

My laboratory's research on sound-source localization (*where* sound sources are) in a real world of moving listeners and sound sources convinced me that to localize *where* a sound source is in the real world, the *brain* must combine information derived from the auditory-spatial cues with that derived from the position of the head (see Pastore et al., 2020; Yost et al., 2021, for reviews). The position of the head must be specified in relationship to the world surrounding the listener and the objects in that world.

These observations and ideas led to what was for me a serendipitous conclusion, that determining *what* a sound source is and *where* it is requires the *entire brain*, not just the *auditory brain*, and almost nothing seemed to be known about how the *entire brain* does either. Clearly, I was not the only one making this observation, so I benefitted from a zeitgeist regarding how little is known about the *entire brain* and auditory perception.

My final bit of serendipity was the realization of how crucial memory is for successfully functioning in an auditory world. Auditory memory is an understudied and, in my opinion, underappreciated aspect of auditory perception. Earlier in this essay, I observed that listeners have little difficulty knowing the source of a sound when a *car* screeched, a *baby* cried, a *ball* bounced, or the *leaves* on a tree rattled. I had the serendipitous realization, however, that the ability to identify the sources of these sounds immediately with little or no effort requires that key features of the sounds and their labels had already been stored in memory and that when a sound occurs, the appropriate labels for its source can be easily recalled from memory. *What* are those auditory features, *where* is this information stored, *how* is the information stored

and then retrieved, *when* in the time frame of auditory processing do these events take place? As far as I know, the answers to these key scientific questions are poorly understood, if they are understood at all.

What about the *brain* and the issue of *where* sound sources are located? I already mentioned our arguments that sound-source localization in the real world requires combining information about the auditory-spatial cues with that from the position of the head (again, see Pastore et al., 2020; Yost et al., 2021, for reviews). The 2014 Nobel Prize in Physiology and Medicine was awarded to John O'Keefe, Eduard Moser, and May-Britt Moser for their work on spatial memory ("Cells that constitute a positioning system in the brain" from their Nobel citation; see bit.ly/4hcGEDK). When I first learned of the award, I did not think their work had much to do with my interests in auditory perception, even though I had read a few of O'Keefe's papers in the early 1980s. It wasn't long until I had a serendipitous moment realizing their work does have a great deal to do with issues of the *entire brain*'s ability to determine *where* sound sources are located.

The Nobel Laureates and others have described in some detail the formation of *cognitive spatial maps* in the hippocampus and surrounding entorhinal cortex (e.g., O'Keefe, 1991). Indeed, *cognitive spatial maps* are what you depend on to navigate in your bedroom at night when you can barely see, and what you can recall about the spatial layout of your bedroom even when you are not in it.

A thought experiment might demonstrate the possible impact of *cognitive spatial maps* for determining *where* sound sources are located. Imagine that you are awakened in the middle of the night in your totally dark bedroom by a brief banging sound. The auditory-spatial cues led you to realize that the sound came from directly in front of you as you lie in your bed, at slightly below the height of you in your bed, and at some distance from you. You immediately conclude, with little effort, that the sound came from the bathroom; perhaps the picture fell off the wall.

How did you conclude that the picture fell off the wall in your bathroom? Your brain would be able to use the auditory-spatial cues to determine that the source of the sound was directly in front of your head as you lay in bed. Your *cognitive spatial map* might indicate that there is a picture on the wall in the bathroom directly in front of you as you lay in bed. Note that the *cognitive spatial map* does not indicate the position of objects (e.g., the picture) just relative to your head, but relative to the room and the other objects in it, a condition that is required to localize the sources of sounds in the real world (O'Keefe, 1991). Several different sensory cues have been shown to generate *cognitive spatial maps*, but it is not clear if auditory cues are used. And, if so, how? Do such maps assist in sound-source localization as the thought experiment suggests? If so, how?

My conclusion from my recent serendipitous episodes is that knowing a lot more about auditory memory seems to be a worthwhile research endeavor, which I would peruse if I wasn't retired. A pursuit, which will probably be advanced by many additional serendipitous moments, which others, not I, will hopefully have.

References

Bregman, A. S. (1990). *Auditory Scene Analysis*: *The Perceptual Organization of Sound.* MIT Press, Cambridge, MA.

Boring, E. G. (1942). *Sensation and Perception in the History of Experimental Psychology*. Appleton-Century-Croft, Inc., New York, NY.

Brownell, W. E. (2017). What is electromotility? The history of its discovery and its relevance to acoustics. *Acoustics Today* 13(1), 20-26.

Lutfi, R. (2007). Sound source identification, In Yost, W. A., Popper, A. N., and Fay, R. R. (Eds.), *Auditory Perception of Sound Sources*. Springer. New York, NY, pp. 13-42.

O'Keefe, J. (1991). An allocentric spatial model for the hippocampal cognitive map. *Hippocampus* 1(3), 230-235.

Pastore, M. T., Pulling, K. R., Chen, C., Yost, W. A., and Dorman, M. F. (2024). Synchronizing automatic gain control in bilateral cochlear implants mitigates dynamic localization deficits introduced by independent bilateral compression. *Ear and Hearing* 20, 1097-1113.

Pastore, M. T., Zhou, Y., and Yost, W. A. (2020). Cross-modal and cognitive processes in sound localization. In Blauert, J., and Braasch, J. (Eds), *The Technology of Binaural Understanding*. Springer Cham, Cham, Switzerland.

Yost, W. A. (1991). Auditory image perception and analysis, *Hearing Research* 56, 8-19.

Yost, W. A., Pastore, M. T., and Zhou, Y. (2021). Sound source localization is a multisystem process. In Litovsky, R., Goupell, M., Popper, A. N., and Fay, R. R. (Eds.) *Binaural Hearing* Springer Cham, Cham, Switzerland.

Yost, W. A., Popper, A. N., and Fay, R. R. (Eds.). (2007). *Auditory Perception of Sound Sources*. Springer, New York, NY.

Contact Information

William A. Yost
william.yost@asu.edu

9119 N. 117th Way
Scottsdale, Arizona 85259, USA