I'm not too sure when cocktail parties started, but the now popular "cocktail" was defined in the early 1800s, so let's assume they've been going on for 200 years or so. And I'm betting that, whether it was gold miners at a martini party in San Francisco in the 1860s, or a group of friends sampling Margarita Sames's exotic concoction at her 1948 Christmas party in Acapulco, people have always had trouble understanding each other at these events.

In fact, "the cocktail party effect" is considered a scientific term, first written about extensively in the 1950s. You know how it goes. When there are multiple talkers and background noise, you attempt—and sometimes are able—to attend to the talker you consider most important or to something else that draws your attention. Audibility, of course, is a key player here, but how we do this is a bit more complex than just "hearing" everything. And what happens when the listener has a hearing loss, distortions are present, and some of the audibility is lost?

To learn all about how we understand speech in difficult listening situations, it seemed only appropriate to bring in an MIT-educated *electrical engineer*. But before you grab your slide rule, understand that this is an engineer whose research includes studies of auditory attention, sound source separation, spatial hearing, cross-modal integration, neural coding, and perceptual plasticity. **Barbara Shinn-Cunningham**, PhD, is Associate Professor of Cognitive and Neural Systems and Associate Professor of Biomedical Engineering at Boston University. She also is an instructor in the Harvard/MIT Health Science and Technology Program and an adjunct professor of the Naval Postgraduate School. She is the recipient of numerous awards, including fellowships from the Alfred P. Sloan Foundation, the Whitaker Foundation, and, most recently, the National Security Science and Engineering Faculty Fellows program.

Barbara tells us that her interest in helping listeners with hearing impairment was sparked a few years ago when her elderly uncle visited at Christmas—and immediately removed his new high-end hearing aids when he sat down at the dinner table. When not working, she tries to get her husband and two teenaged sons interested in auditory perception by making them listen to her play the oboe. Fortunately, she put down her oboe just long enough tell us all about this so-called cocktail party effect, and most importantly, how it impacts our patients. Cheers.

GUS MUELLER
Page Ten Editor

"I want to party, but my hearing aids won't let me!"

By Barbara Shinn-Cunningham


Barbara Shinn-Cunningham

1 I hear that you're doing research on hearing aid use and benefit. Is this true?

Well, not exactly... but I *want* to!

My background is in electrical engineering, but for the last 25 years I have been studying hearing and auditory perception. My research has been very basic, though, working to understand how normal-hearing, young adults process sound.

I started out studying how we can tell the location of a sound source, and gradually got interested in why "spatial hearing" really seems to matter a lot in everyday settings, like at a cocktail party. The more I looked into these issues, the more I realized that understanding how normal-hearing listeners make sense of sounds at a cocktail party can lead to insights that could allow us to build better hearing aids.[1] Lately, I have started studying what happens in "cocktail party" situations to hearing-impaired listeners.[2]

2 Great! Maybe you can tell me why so many of my patients complain that their hearing aids are useless in social settings, like at a party. What's going on?

In active social settings, such as a busy restaurant, there are lots of separate similar-sounding sources coming into the ears from all directions. Normal-hearing, young, healthy listeners are good at focusing on whatever source they're interested in (like their witty co-worker across the table) and ignoring other sounds (the snob opining about the hint of grapefruit in his chardonnay, the couple bickering over the check...). In contrast, hearing-impaired listeners have real trouble filtering out unwanted sound sources, a process known as *selective attention*.

3 But I know that I fit my patients' hearing aids well. All those competing sources should all be audible! Isn't that good enough?

Unfortunately, no. Lack of audibility is not the only problem that interferes with understanding sound sources in real-world settings. For instance, imagine two talkers a few feet away from you. Normal speech comes in concentrated bursts. It is "sparse" in time and frequency. As a result...

4 Wait a minute... What do you mean by "sparse"?

Well, think about a typical sentence. There are a lot of silences and gaps between the syllables, words, and sentences. In this sense, speech is "sparse" in time.

Frequency is really important, too. The ear breaks down sound into constituent frequencies (high down to low), so when sound enters the brain, each frequency is represented in its own "channel." Different sets of neurons represent the sound in each narrow range of frequencies. This is important, because if you look at the frequency content of speech, it usually has energy only at some frequencies at one instant in time. Thus, speech is "sparse" in frequency.

5 Okay, got it. But how does this relate to understanding?

The fact that speech is sparse in time and frequency is a really useful thing when it comes to listening to speech at a party. It is fairly likely that if there are two talkers speaking at the same time, there will be moments when the speech of one talker has a gap, allowing you to get a clean "glimpse" of what the other talker is saying.[3]

Moreover, speech is also sparse in frequency. If the frequency content of two sounds doesn't overlap, the inner ear can represent *both* sounds without their interfering with one another, *even if* they are both on at the same time. This gives you an even better chance of being able to get "glimpses" of what one talker is saying in the presence of another talker.

If you have normal hearing, the fact that natural speech is sparse means you are often *capable* of hearing all of the sound coming from *each* of two simultaneous talkers if both are audible. That is really helpful when it comes to communicating at a party. But, the problem is that even when both talkers are audible, that doesn't guarantee you can ignore one and pay attention to the other.

6 Can you give me an example of what you mean?

Imagine that you are trying to listen to Tom and his identical twin brother, Jerry, is standing next to him talking on his cell phone. It is easy to get confused and have trouble focusing on what Tom says. That both of the competing sources are audible doesn't guarantee you can selectively attend to whichever one you want.

When a hearing aid wearer complains that his hearing aids "just amplify all of the background," it suggests that lots of sounds are audible; the problem is that the user cannot focus attention on a desired source and filter out the other sources.

7 Oh, I see. Back to your baffling "selective attention," right?

Yes, indeed. Selective attention is the process that allows the brain to throw away unimportant information and concentrate on the important stuff. The brain cannot simultaneously process all the sounds present in the world—there is just too much going on. This is especially true in a social setting with lots of people who are talking, laughing, and interacting. Instead, we typically process and understand one sound source at a time. The decision about what to listen to at a given moment is controlled by "executive processes" in central portions of the brain. These executive processes determine what we attend and what we ignore at any given moment.

8 Are you trying to tell me that a peripheral hearing impairment causes problems with some mysterious "executive processes" in the brain?

No, no! Not at all! It is true that aging may interfere with executive function, and I know that many of your patients are older. Your older patients *may* have trouble with executive function, which will exacerbate any problems they have due to hearing loss.

But even otherwise-healthy, young listeners with peripheral hearing impairment have trouble with selective auditory attention. The problem is not that the executive control center in the brain isn't working well, it's that a peripheral hearing impairment distorts sound in ways that make it difficult to separate an important sound source from competing sources. If your brain doesn't correctly separate the sound coming from the person you want to hear from all the other sounds around you, selective attention will fail, whether or not the more central processes are functioning properly.

9 What do you mean by "correctly separate?"

When healthy, young listeners hear a mixture of sound sources, their brains naturally and automatically begin to sort the sound into constituent sources, divvying up the mixture into distinct "objects" (perceptual estimates of the sound coming from each different physical sound source). This is known as "auditory scene analysis" or ASA. ASA is such an important issue in auditory perception that Al Bregman wrote a whole book on the topic in the early 90s.[4] Figuring out how normal-hearing listeners form perceptual auditory objects from a mixture of sound continues to be an active area of scientific research (and one we've been pursuing in my laboratory).

Imagine yourself in your kitchen talking to your son about his math test while you sauté some chicken. The phone rings, and when your son turns to get it, he knocks a glass off the counter and it shatters on the floor. Each of those physical objects—your son's voice, the sizzling meat, the phone ringing, the exploding glass—creates sound that adds acoustically to all the other sounds. But you perceive each one as a separate thing—an object. For instance, you don't hear the shattering glass and your son's voice as coming from the same source. For a normal-hearing listener, most everyday sounds are naturally perceived as separate objects. However, accomplishing this is an amazing feat—one that is disrupted by hearing loss.

10 It seems obvious that we would hear each sound source as a different thing. Are you saying it is actually hard?

Yes, it is hard, in the sense that it is nearly impossible to write down a complete set of rules describing how to take a mixture of sound and properly separate it into distinct perceptual objects. Yet, most healthy young listeners do this effortlessly. In fact, even though lots of researchers are trying to build machines that can properly separate a sound mixture into objects, so far none comes close to solving this problem as well as a human listener does.

Research *has* identified some of the acoustic cues that signal to our brains what sound should be grouped together into one perceptual object (what sound probably came from the same physical source). And most of these cues work in an intuitive way. For instance, if two sounds start and stop at the same time, we are likely to hear them as part of the same object. That makes a lot of sense, because two independent, physical sound sources usually wouldn't start and stop together. (In fact, if distinct physical sound sources start and stop together, it usually isn't by chance! For instance, while playing my oboe at my orchestra rehearsal tonight, I had to concentrate on the conductor to ensure that the flute and I played together!)

11 I get the basic idea. But what are the cues that allow normal-hearing listeners to separate sound into objects?

Well, when sounds share common fluctuations in intensity (getting louder and softer at the same time) we are more likely to group them together. We tend to group together sound that is coming from the same direction or shares a common pitch (the attribute of musical notes that conveys melody). Sound that changes smoothly over time and that sounds "similar" from one moment to the next (like my voice when I am speaking) causes sound to group into an object.

Basically, when sound elements share attributes that suggest that the same physical process generated them (as when they turn on and off simultaneously), our brains tend to group those elements together. This allows us to separate a mixture of sounds into different objects.[5,6]

12 So, how does separating sound into objects relate to selective attention?

Our central executive processes select one object at a time from all the things going on around us. When we focus attention, we really *hear* one object—at the cost of not *really* hearing the competing objects (even though they are audible). In other words, perceptual objects are the "units" on which attention operates. For selective attention to operate effectively, we have to be able to properly separate the sound mixture into objects, so that we can selectively focus attention on the one we want to understand.[7]

In the Tom and Jerry example, if your brain fails to separate the brothers' speech into distinct objects, you will hear a jumbled combination of words from both brothers even if you try to focus attention on one. This will make it hard to understand the meaning of either since the object you hear will be a mixture of the two voices. Imagine what would happen if instead of selectively hearing either "Tom" or "Jerry," you heard the nonsensical "ToJerrmy."

13 But as long as we perceive distinct objects, can we selectively attend to whatever we want?

Well, most of the time. As long as the sound sources are perceived as distinct objects, we typically attend to one source at a time.

But which object you focus on depends not only on what you are *trying* to attend, but also on what sounds are out there in the world. Imagine yourself back in the kitchen. The sound of the exploding glass was probably loud, sudden, and unexpected. This makes it very salient—and salient sounds can seize your attention even if you aren't *trying* to listen to them.

This is true not only in hearing, but in other sensory modalities as well. A sudden flash of a bright light can "grab" your attention, even if you aren't trying to look at it. In fact, new hearing aid users often are distracted by environmental sounds that they haven't heard for years, because they are novel (and therefore salient). At any given moment there can be a quiet battle going on between what you are volitionally *trying* to pay attention to and a salient object in the scene.

Scientists describe this battle as a "biased competition" for your attention.[8] Objects in a scene all compete for your attention. However, your desire to selectively attend to one particular thing gives that object a leg up, biasing the competition so that the selected object is more likely to "win" and become the focus of attention.

It may seem counterintuitive at first, but having salient objects able to win out sometimes over what you *think* you want to attend actually is a very sensible way for the brain to work. Even if I am concentrating on my co-worker's description of the latest antics of our boss, I don't want to miss hearing the fire alarm.

14 Hold on. You make it sound as if I always attend to only one thing at a time. What about when I listen to my spouse and the TV at the same time?

The thing is, you may *think* you're listening to two things at once, but you probably aren't.

Speech is redundant. You can miss a word or a syllab__ here or there and still underst__ a sentence's meaning (here you probably had little trouble filling in "syllable" and "understand"). You can sneak peeks between the TV show and your wife's fascinating story, and still get *most* of each. But imagine getting really engrossed in a particularly exciting car chase... and then suddenly realizing your wife is mad because you didn't switch back in time!

In other words, when there are multiple things going on at the same time, we switch attention between them. Often this works well enough that we *believe* that we are "multi-tasking" when we are really just switching between tasks. In fact, there are measurable costs to switching attention between objects—both because it actually takes time to switch attention from one object to another, and because you become better at focusing attention on a selected object and filtering out competing sounds as you maintain attention on one thing.[9]

15 This all helps explain why my Aunt Bertha has more trouble at a cocktail party than when she's talking to Uncle Byron and me over dinner.

Yes. The more "social" a gathering, the more sound sources there are. The more sound sources there are, the more they will overlap with one another in time and frequency, making it harder to segregate the mixture properly into objects. And without proper object formation, selective attention is not effective.

On top of that, the more people there are in a conversation, the more often the conversation jumps around, unpredictably, from one talker to another. Whenever the conversation jumps, the listener has to switch attention and refocus on the new talker, which takes time. Moreover, you have to keep up with the rapid shifts in the conversation or you will be left out and feel socially isolated. To really participate, you must be able to efficiently pick out the right object to attend at each moment.

16 How do people usually select which object to listen to?

Well, you can focus your attention in different directions (to your right, above you); spatially directed attention is very powerful and (pardon the pun) has received a lot of attention. But you can also direct attention to a distinct pitch (the high voice), intensity (the softer talker), voice quality (the woman amidst the men), or even expected rhythm and content (like listening for a phone number in the middle of a rowdy party). What this means, though, is that the more objects there are, the more likely it is that the one you want to attend is a lot like an object you are trying to ignore.

17 Okay, but let's get back to my original question. Exactly *why* do hearing-impaired listeners have trouble with selective attention?

Although there are a whole host of reasons why hearing impairment interferes with understanding sound in a complex setting, one big factor undoubtedly is that hearing-impaired listeners have trouble properly separating a sound mixture into objects.

Hearing impairment degrades the representation of sound in the brain in a lot of ways. First of all, the different "frequency channels" that I mentioned before get broader with hearing impairment. Consequently, competing sound sources are effectively less "sparse" in their neural representation. Simultaneous sound sources that could be represented in separate neural populations in a healthy ear are very likely to overlap with one another and cause perceptual interference in a hearing-impaired ear. This makes it harder to separate the brain's response to an unattended sound source from the response to the attended sound source. In the hearing-impaired ear, sound sources are less likely to be perceptually separable than in a healthy ear.

On top of the loss of frequency specificity in the impaired ear, temporal structure is often degraded. Hearing-impaired listeners are less able to detect temporal gaps in stimuli and less able to use things like harmonicity and pitch. Hearing impairment can also degrade spatial perception. The list goes on and on. The very cues that enable normal-hearing listeners to segregate the sound sources in a mixture into distinct perceptual objects are degraded with hearing impairment.

18 Great. So what can I do for my patients so they can participate in everyday social settings?

Unfortunately, you probably can't do a whole lot more than what you are already doing. It turns out that the acoustic attributes that convey the meaning of speech are the same ones that allow listeners to segregate sound sources in an auditory scene: changes in amplitude and frequency over time, temporal structure, etc. A lot of great engineers have worked hard to optimize current hearing aids to best convey speech meaning. As a result, today's hearing instruments are probably close to being optimized to convey the features of sound that allow listeners to analyze the auditory scene.

The real problem is that hearing aids cannot bypass the impaired ear. The impaired ear just doesn't have the same temporal or frequency resolution as a healthy ear—resolution that is necessary to efficiently separate competing sound sources.

19 But what about directional hearing aids? Can't they help suppress competing sources?

Indeed, my colleagues tell me that directional hearing aids sometimes help listeners participate in a social setting. Directional aids may simplify the scene enough that listeners can *at least* understand the source in front of them, which is better than not being able to understand *any* of the sources in the scene. However, in everyday settings, they actually don't provide that much benefit.

20 So then, what is your vision for the future?

We have to build smarter hearing aids. We need hearing aids that present simple enough auditory scenes that even an impaired ear can analyze them. We need to devise algorithms that automatically analyze the auditory scene into objects. But then we need to choose *which* objects to present (and *how many* to present) intelligently.

Specifically, the choice of what to present should depend not only on the auditory scene itself and what objects are in that scene, but also on what matters to the listener at a given moment. Although it may be years from becoming feasible, I believe this is the only path toward enabling hearing-impaired listeners to participate more actively at a cocktail party and other social events.

REFERENCES

1. Shinn-Cunningham BG, Best V: Selective attention in normal and impaired hearing. *Trends Amplif* 2008;12(4):283-299.
2. Best V, Marrone N, Mason CR, et al.: Effects of sensorineural hearing loss on visually guided attention in a multitalker environment. *J Assn Res Otolaryngol* 2008, pubished online Nov. 14.
3. Cooke M: A glimpsing model of speech perception in noise. *J Acoust Soc Am* 2006;119(3):1562-1573.
4. Bregman AS: *Auditory Scene Analysis: The Perceptual Organization of Sound.* Cambridge, MA: MIT Press, 1990.
5. Carlyon RP: How the brain separates sounds. *Trends Cogn Sci* 2004;8(10):465-471.
6. Darwin CJ, Carlyon RP: Auditory grouping. In Moore BCJ, ed., *Hearing*. San Diego: Academic Press, 1995: 387-424.
7. Shinn-Cunningham BG:Object-based auditory and visual attention. *Trends Cogn Sci* 2008;12(5):182-186.
8. Desimone R, Duncan J: Neural mechanisms of selective visual attention. *Ann Rev Neurosci* 1995;18:193-222.
9. Best V, Ozmeral EJ, Kopco N, Shinn-Cunningham BG: Object continuity enhances selective auditory attention. *Proc Nat Acad Sci* 2008;105(35):13174-13178.