

# Speech: It's Not as Acoustic as You Think

**Matthew B. Winn**

*Address:*

Department of Speech &  
Hearing Sciences  
University of Washington  
1417 NE 42nd Street  
Seattle, Washington 98105  
USA

*Email:*

mwinn83@gmail.com

*Address after August 2018:*

Speech-Language-Hearing  
Sciences Department  
University of Minnesota  
164 Pillsbury Drive SE  
Minneapolis, Minnesota 55455  
USA

*Email:*

mwinn@umn.edu

*There are many ways that we understand speech by paying attention to things other than what is spoken.*

*The waveforms patterns bounced up and down as if to evade my grasp. Spectrograms, supposedly there to help me quantify the exact difference between one vowel and the next, instead seemed to mock me with their peculiar and elusive patterns, changing with every utterance and every talker. Was it possible to really crack the speech code? The deeper I looked, at each frequency and each millisecond of speech, the farther I felt from understanding actual speech communication. My head recoiling in frustration, I sighed and turned away from the computer screen. As it turned out, looking away from the acoustic signal was a step in the right direction.*

There is a humbling paradox of studying speech acoustics, which is that speech communication seems so rapid, effortless, and universal, yet scientific inquiry into the topic is overflowing with challenges and a vast ocean of unknowns (for a further discussion of this, see the article by de Jong in this issue of *Acoustics Today*). The history of speech science is full of investigators tortured by the exhaustive search for the understanding of speech based on the acoustic signal itself. What could be more satisfying than to decode the key elemental sounds that make up the way humans communicate, absorb news, greet friends, tell jokes, and express thoughts? Best of all, the science of speech acoustics would provide objectivity. The utterance is familiar, and it can be measured in numerous ways. There is no need for subjective or enigmatic forays into the world of linguistic structure, hazy impressions of what is proper or improper, or indeterminate variations in meaning and intention. If humans speak the sound, it should be possible to decode it with meticulous attention to the measurements. There is a kind of comfort in the nitty-gritty, irrefutable measurements of sound pressure, frequency, and duration.

But it is possible to let this deep pursuit temporarily obstruct the view of the full expanse of the information used to understand speech, and there is also a risk of having too narrow a scope of what information should be regarded as essential to the communication process. With that in mind, this article is about how spoken language is not as acoustic as one might think. The number of ways that we take in and use information to interpret speech, apart from hearing the acoustic details of the signal, is staggering, and inspiring, especially to this author who identifies as an auditory scientist. Some examples are unsurprising, but others might stun you. The hope is that we can search for a more complete understanding of speech by paying special attention to the parts that are actually unspoken.

To begin with, it should be noted that there is a good deal of useful information to be gleaned from the speech signal itself. A listener succeeds most when the frequency range of speech is audible and when there is sufficient perception of

both frequency and timing information. Hearing thresholds, combined with factors related to aging, can certainly explain a substantial amount of performance observed in the laboratory and clinic (Humes, 2002; Akeroyd, 2008). Toward an understanding of the essential speech components necessary for the transmission of speech across telephone lines, Fletcher and Galt (1950) divided the sound spectrum into frequency bands and measured the contribution of each band to intelligibility. Deconstructing the signal into its most essential components is reminiscent of other sciences and has proven valuable. Similar approaches were taken by Miller and Nicely (1955) and revisited by Allen (1995) to understand how the energy in the signal is recovered and used by the listener in very specific ways. These principles are still used today in audiology clinics where the audibility of different speech frequencies are used to estimate the performance of a person being fit with hearing aids.

Despite the intuitiveness and utility of explaining speech communication based solely on acoustic audibility, we get more information than acoustics alone provides. Some reports claim that the variance in ability to perceive speech is not only affected but also *mainly* affected by nonauditory abilities (George et al., 2007). Other studies place the contribution of nonauditory information at 30% to 50% (Humes, 2007). In an influential report collecting observations across many previous studies, Massaro and Cohen (1983) suggested that we embrace the integration rather than force the separation of auditory and nonauditory streams. As a consequence, it is clear that one must give attention to contextual factors when examining the acoustic properties of the speech signal.

### Talker Familiarity

It will come as no surprise that it is easier to understand a talker who speaks your particular dialect (Labov and Ash, 1997), and the benefits scale to the relative distance of dialect from your own (Wright and Souza, 2012; also see the article by de Jong in this issue of *Acoustics Today*). Furthermore, there are special benefits when listening to a longtime partner or spouse whose speech is measurably more intelligible to the partner than to strangers (Souza et al., 2013). Talker familiarity can yield benefits even on very short timescales because sentences are more intelligible when preceded by sentences spoken by the same talker (Nygaard et al., 1995). Notably, the counterbalancing of talkers and listeners in these studies demonstrates that acoustic factors cannot explain the effect. Instead, some other property of the listener or the relationship history, such as knowing a person's

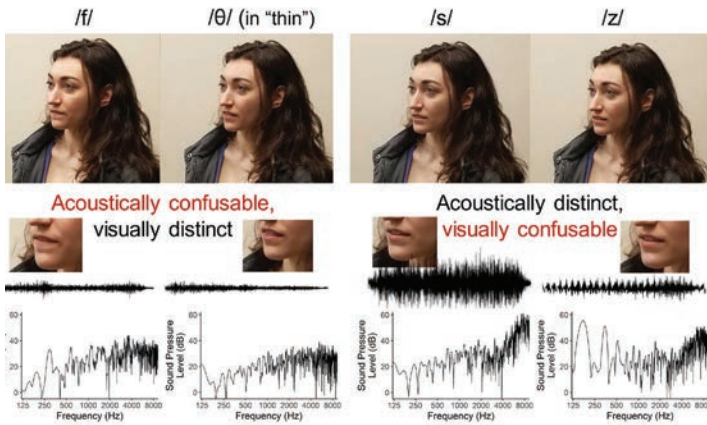
dialect, the funny way “water” is said, or the words typically used, can make a substantial difference.

Apart from the intuitive advantage of knowing who's talking, there are some rather unusual and surprising effects of context. Simply changing the *expectation* of what a talker sounds like can affect how speech is perceived, even if the acoustics have remained unchanged. Intelligibility is poorer when utterances are *thought to* be produced by a person who is not a native speaker of one's own language, even if the sound has not been changed (Babel and Russell, 2015). We routinely accommodate the acoustic difference between a woman's and a man's voice by, for example, expecting higher or lower frequency sounds, and this accommodation can be induced if the listener simply *sees* the talker's face (Strand and Johnson, 1996), especially if the listener has a hearing impairment (Winn et al., 2013).

Indeed, even seemingly unrelated objects in the environment, like stuffed animals, can affect speech perception! Hay and Drager (2010) conducted a clever vowel-perception experiment that hinged on a listener's knowledge of the differences in dialect between Australian and New Zealand varieties of English. Imagine a vowel sound that is intermediate between the sounds that Americans would use for the vowels in the words “head” and “hid”; it would be a word that could fall either way. This vowel sound would be heard in the word “hid” in Australia but be perceived as “head” in New Zealand. In an experiment where stuffed toys were placed conspicuously in the room, the listeners were more likely to hear this ambiguous vowel as Australian “hid” when the toy was a kangaroo but were biased toward hearing “head” (a New Zealand interpretation) when the toy was a kiwi bird. The toy primed the listener to match the sound to the appropriate dialect that would be associated with the animal despite no difference in the acoustic stimulus.

### Visual Cues

Visual cues are used by all sighted listeners, not just those with hearing loss. The complementarity of the ears and eyes in speech perception is remarkable. Speech sounds that are most acoustically similar (think of “f” and “th” or “m” and “n”) are reliably distinguished visually (see **Figure 1**). The reverse is true as well. Even though it is nearly impossible to see the difference between “s” and “z” sounds, listeners almost never mistake this phonetic contrast (called *voicing*, referring to glottal vibration felt in the throat) even when there is significant background noise (Miller and Nicely, 1955). The learned association between the sounds and



**Figure 1.** Speech sounds that are acoustically similar (like the first consonants in the words “*fn*” and “*thin*”) are visually very distinct. Sounds that are visually similar (like “*s*” and “*z*”) are acoustically quite distinct, as seen in the stark differences in the low-frequency region of the spectra (**bottom row**).

sights of speech is acquired in infancy (Kuhl and Meltzoff, 1982), including the ability to detect whether a silent talking face is speaking the emergent native language of the infant observer (Weikum et al., 2007).

Visual cues are so powerful that it can be hard to suppress them even when they are known to be phony. The most well-known example of this was made famous by a series of experiments by McGurk and McDonald (1976) in which the audio and video tracks of a single syllable were mismatched (e.g., the audio of /pa/ combined with the video of /ka/). In this situation, the visual stream will “fuse” with the auditory stream, yielding perception of /ta/, which is intermediate between the two source signals, even if the listener knows she is being tricked. This phenomenon has come to be known as the “McGurk effect,” and there are many different variations that have emerged in the literature (with many examples available online on YouTube, e.g., [acousticstoday.org/mcgurk](http://acousticstoday.org/mcgurk) and [acousticstoday.org/mcgurk2](http://acousticstoday.org/mcgurk2)). Even when the video is clearly from another talker (including mismatching stimuli from women and men; Green et al., 1991), the effect can be hard to suppress. There is evidence to suggest that auditory-visual fusion can be somewhat weaker in languages other than English (Sekiyama and Tohkura, 1993) but for reasons that are not totally understood. Auditory-visual fusion is, however, a much stronger effect for people who have a hearing impairment (Walden et al., 1990), consistent with the relatively greater reliance on lip reading. There are also reports of speech sounds that are affected by *touching* the face of a talker as well (Fowler and Dekle, 1991) which,

although not a typical situation, demonstrates that we are sensitive to multiple kinds of information when perceiving speech.

### Linguistic Knowledge and Closure

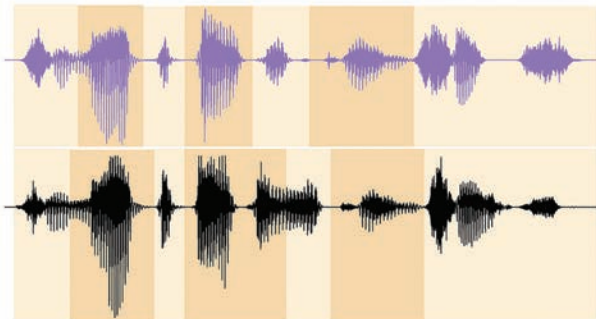
The same speech sound can be heard and recognized differently depending on its surrounding context. Words that are spoken just before or just after a speech sound help us interpret words in a sensible way, especially when the acoustic signal is unclear. Consider how the word “sheets” is entirely unpredictable in the sentence “Ashley thought about the sheets.” Ashley could have been thinking about anything, and there’s no reason to predict the word sheets. Conversely, if there is some extra context for that word, such as “She made the bed with clean...,” then the listener’s accuracy for recognizing “sheets” becomes *more accurate*, presumably because of the high amount of context available to help figure out what word fits in that spot. That context-related benefit can work both forward (prediction of the word based on previous cues) or backward (recovery of the word based on later information) in time. In either case, we are driven to conclude that the acoustics of the target word “sheets” can be studied to the level of the finest detail, yet we cannot fully explain our pattern of perception; we must also recognize the role of prediction and inference. Taking fragments of perception and transforming them into meaningful perceptions can be called “perceptual closure,” consistent with early accounts of Gestalt psychology.

There was once a patient with hearing loss who came into the laboratory and repeated the “... clean sheets” sentence back as “She made the bagel with cream cheese.” This was an understandable mistake based on the acoustics, as shown in **Figure 2**; the amplitude envelopes of these sentences are a close match. If we were to keep track of individual errors, it would seem that “bed” became “bagel,” “clean” became “cream,” and “sheets” became “cheese.” These are errors that make sense phonetically and acoustically. However, it is rather unlikely that the listener actually misperceived *each* of these three words. It is possible instead that once “bed” became “bagel,” the rest of the ideas simply propagated forward to override whatever the following words were. **Figure 2** illustrates the mental activity that might be involved in making these kinds of corrections.

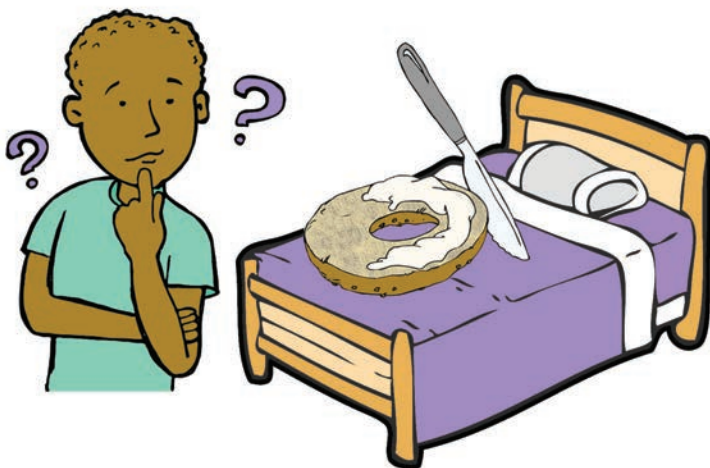
A clever example of the listener’s knowledge overriding the acoustics was published by Herman and Pisoni (2003) who replaced several consonants in a sentence with different con-

## Speech: Not So Acoustic

“She made the bed with clean sheets”



“She made the bagel with cream cheese”



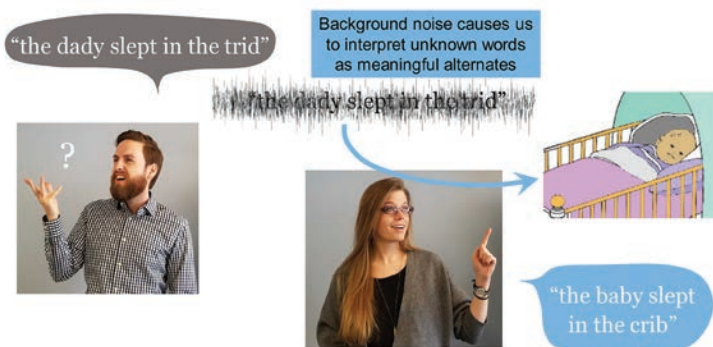
**Figure 2.** Making the bed “with cream cheese” doesn’t make sense... it must have been “with clean sheets.” The amplitude envelopes of the sounds are nearly identical (top), which might explain why a seemingly drastic series of mistakes could be made. Within each interpretation, the words make sense in relation to each other; the listener would probably not think of “bagels” and “clean sheets” as belonging in the same sentence.

sonants articulated at the wrong place in the mouth (e.g., “His plan meant taking a big risk” transformed into “His tlan neant tating a did rist” (see **Multimedia File 1**, parts **a** through **d**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic)).

They called the mispronounced utterances “elliptical speech.” Although these sentences were very challenging to hear in quiet, a little bit of background noise invoked the strategy of using “top-down” processing to replace each misarticulated work with a more reasonable counterpart. The conversion of a non-sensical utterance into a meaningful one is shown in **Figure 3**.

### Even the Acoustic Parts Are Not as Clear One Thinks

It seems intuitive to think of speech as sentences constructed from smaller units like words, which are themselves con-



**Figure 3.** An utterance like “the dady slept in the trid” is non-sense, but if the same utterance is heard with sufficient background noise, the brain replaces the unusual words with similar-sounding alternates that create a well-formed sentence like “the baby slept in the crib.” See also **Multimedia File 1**, parts **a** through **d**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic).

“His shoes were untied”

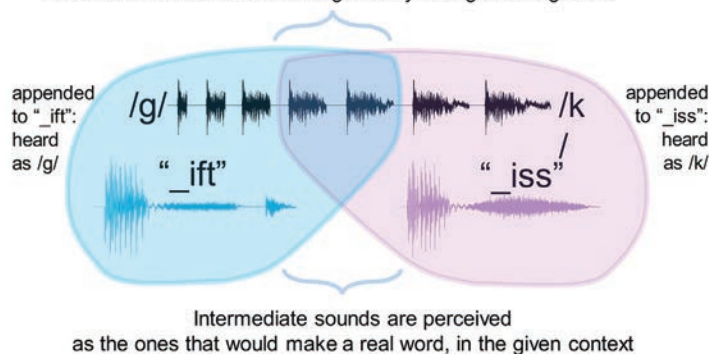


Listener can unpack this ambiguous utterance into its components “his” + “shoes”

**Figure 4.** Articulatory motions for speech sounds overlap in time, producing utterances that would be confusing if heard out of context. Within the context of this sentence, “hish” can be interpreted as “his” + the onset of “shoes” (see **Multimedia File 2**, parts **a** and **b**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic)).

structed from smaller units called phonemes. This building block framework not only fails to account for the influence of the aforementioned “top-down” factors, it also assumes that the smaller building blocks (individual phonemes and words) are themselves intelligible. However, this is frequently not the case. If the words from a sentences are stripped out and played in isolation, they are sometimes entirely unintelligible, again highlighting the importance of the surrounding context when identifying what is spoken. The intelligibility of short utterances excerpted from a conversation was analyzed by Pickett and Pollack (1963), who found that performance was only above 50% when substantial surrounding context to a word was provided. **Figure 4** shows how this is true for the sentence “His shoes were untied.” The first word in isolation sounds like “hish,” not like any recognizable English word (see **Multimedia File 2**, parts **a** and **b**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic)). But when followed by the rest of the sentence, the original meaning

Continuum of consonants that gradually change from /g/ to /k/



**Figure 5.** The “Ganong effect.” When labeling a consonant that is morphed between /g/ and /k/, there will be more perceptions of /g/ when the consonant is followed by a syllable like “ift” because “ift” is a real word. If the same sound is followed by “iss” then more items in the continuum will be labeled as /k/. Perception is driven not just by acoustics but also by the lexicon.

can be recovered. The same is true for almost any casually spoken sentence. In “I’ll talk to you later,” the final four syllables usually come out as a blur with vowels replaced with mumbles (see **Multimedia File 3**, parts **a** and **b**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic)), and in “This is as good as it gets,” some of the first few vowels are dropped entirely, leaving a sequence of buzzing “z” sounds for the listener to unpack as whole words (see **Multimedia File 4**, parts **a** and **b**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic)).

### Filling in the Gaps

The influence of nonacoustic factors like linguistic knowledge is especially noticeable, and helpful, when the speech signal is hard to hear. When individual words are completely masked by noise, a listener can still fill in the gaps to correctly guess what was spoken; this has come to be known as *perceptual restoration* (Warren, 1970). Not only is perception of the speech smooth and continuous, the listener appears to mostly discard the perceptual details of the noise; if a coughing sound is made while a person hears a sentence, the timing of that cough cannot be reliably judged. In fact, people tend to estimate that it occurs at a linguistically relevant landmark (e.g., where you might put a comma or break between clauses in a sentence), even if its position was truly random. This observation is merely a seed of a larger pattern where people shape their perception of an utterance to match their framework of language.

### “When You Hear Hooves, You Expect Horses, Not Zebras”

Some words are spoken more frequently than others, and it comes as no surprise that they are more easily recognized, by both native and nonnative listeners (Bradlow and Pisoni, 1999). Listeners expect to hear words that have *meaning*. Consider a spoken utterance where the first consonant is ambiguous; was that first sound a /g/ or a /k/? If the mystery sound is followed by “iss,” then you are likely to hear it as a /k/, because “kiss” is a real word and “giss” is not. However, if the *exact same sound* is placed before “ift,” then you are more likely to hear it as a /g/, for just the same reason. This pattern is illustrated in **Figure 5** and can be heard in (see **Multimedia File 5**, parts **a** through **g**, at [acousticstoday.org/speech-not-acoustic](http://acousticstoday.org/speech-not-acoustic)). These scenarios again highlight how nonacoustic factors can play a role when we perceive speech. The acoustics cannot be said to solely drive the perception because the same consonant sound was used in each case. Only the surrounding context and intuition about likely word meaning can explain the bias. This particular lexical bias effect, commonly known as the “Ganong effect” for the author of a

study (Ganong, 1980) who first described it, is observable mainly when the speech sound is ambiguous and not so much when it is spoken normally. However, as discussed below, the underlying processes related to using context might be active earlier and more frequently than what shows up in basic behavioral tests.

The pattern of biasing toward hearing things that are frequent or meaningful extends down to the level of individual speech sounds, which can come in sequences that are more or less common. Listeners can capitalize on likelihood and structure of words in their language to shape what they think they hear. For example, there are more words than end in the “eed” sound than the “eeg” sound, so if a listener is unsure when the final consonant was a /d/ or a /g/ but was certain that the preceding vowel was “ee” (transcribed phonetically as /i/), then the listener could be more likely to guess that the final sound was a /d/.

It can be tempting to interpret nonacoustic effects like perceptual restoration or lexical bias as a post hoc refinement of laboratory behavior rather than a real perceptual phenomenon. However, in addition to using linguistic knowledge to fill in gaps, people also appear to make *predictions* about upcoming speech that might or might not even be spoken. When the audio of a sentence is cut short before the end, listeners can still reliably say what will be spoken (Grosjean, 1980), and can shadow a talker’s voice at speeds that seem impossible to explain by hearing alone (Marslen-Wilson, 1973), suggesting that they quickly generate predictions about what will be heard *before* the sounds arrive at the ear.

One of the most interesting ways to learn about speech and language perception is through the use of eye-tracking studies that follow a listener’s gaze as the person listens. Such experiments demonstrate that listeners rapidly and incrementally perceive the speech signal and act on even the most

subtle bits of information as soon as the sound arrives at the ear. For example, when hearing “The man has drunk the ...,” an observer will look at an empty glass but instead will look at a full glass when hearing “the man will drink the...” (Altmann and Kamide, 2007). These actions make sense when you stop and think about the meaning of the words, but no such deliberate reckoning is needed because the eye gaze shifts are rapid and occur much earlier than the subsequent words like “water” or whatever else should finish the utterance. This is yet another example of how perception of a word (“water”) is not just about receiving the acoustic pressure waves but can also be shaped dramatically by factors that have nothing to do with the acoustics of the word itself.

Prediction of upcoming words is not merely a neat trick that emerges in the laboratory; it is the foundational principle of entire frameworks of speech perception theories (see Lupyan and Clark, 2015). It is why individuals can complete our friends’ and spouses’ sentences and why one can expect a big play when a sports announcer’s voice begins to swell with excitement. Brain-imaging studies have validated the idea of speech perception as a process of continual prediction and error correction rather than a straightforward encoding of the incoming signal. Skipper (2014) shows that there is actually metabolic *savings* afforded by the use of context, contrary to the idea that computing context is a costly extra processing layer on top of auditory sensation. He goes on to say that the majority of what we “hear” during real-world conversation might come not from our ears but from our brain.

### Conclusion

The study of speech acoustics has demanded creativity and collaboration among a variety of experts spanning multiple fields of study, including acoustics and beyond. There is so much literature on the topic that it is easy to lose track of a reality that is perhaps more obvious to a person who does *not* study speech communication: speech is not nearly as *acoustic* as one might think. Speech has been and will continue to be driven in large part by studies of the sounds of the vocal tract and the auditory-perceptual mechanisms in the ear that encode those sounds. It is undeniable that the quality of the speech signal itself plays a large role in our perception; just ask anyone who has hearing loss. However, by recognizing the nonacoustic factors involved in speech perception, one might better understand why computers don’t recognize speech as well as humans; despite hyperspeed detailed analysis of the acoustic signal, only part of the information is in the signal, and the rest lies elsewhere, either in the environment, on the face of the talker, in the statistics of the language, or, more likely, in the mind of the listener.

### References

- Akeroyd, M. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology* 47(Suppl. 2), S53-S71.
- Allen, J. (1995). Consonant recognition and the articulation index. *The Journal of the Acoustical Society of America* 117, 2212-2223.
- Altmann, G., and Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language* 57, 502-518.
- Babel, M., and Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America* 137, 2823-2833.
- Bradlow, A., and Pisoni, D. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America* 106, 2074-2085.
- Fletcher, H., and Galt, R. (1950). Perception of speech and its relation to telephony. *The Journal of the Acoustical Society of America* 22, 89-151.
- Fowler, C., and Dekle, D. (1991). Listening with eye and hand: Cross-modal contributions to speech perception. *Haskins Laboratory Status Report on Speech Research* SR-107/108, 63-80.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6, 110-125.
- George, E., Zekveld, A., Kramer, S., Goverts, T., Festen, J., and Houtgast, T. (2007). Auditory and nonauditory factors affecting speech reception in noise by older listeners. *The Journal of the Acoustical Society of America* 121, 2362-2375.
- Green, K., Kuhl, P., Meltzoff, A., and Stevens, E. (1991). Integrating speech information across talkers, gender, and sensory modality: Female faces and male voices in the McGurk effect. *Perception & Psychophysics* 50, 524-536.
- Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics* 28, 267-283.
- Hay, J., and Drager, K. (2010). Stuffed toys and speech perception. *Linguistics* 48, 865-892.
- Herman, R., and Pisoni, D. (2003). Perception of “elliptical speech” following cochlear implantation: Use of broad phonetic categories in speech perception. *Volta Review* 102, 321-347.
- Humes, L. E. (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *The Journal of the Acoustical Society of America* 112, 1112-1132.
- Humes, L. E. (2007). The contributions of audibility and cognitive factors to the benefit provided by amplified speech to older adults. *Journal of the American Academy of Audiology* 18, 590-603.
- Kuhl, P., and Meltzoff, A. (1982). The bimodal perception of speech in infancy. *Science* 218, 1138-1141.
- Labov, W., and Ash, S. (1997). Understanding Birmingham. In C. Bernstein, T. Nunnally, and R. Sabino (Eds.), *Language Variety in the South Revisited*. University of Alabama Press, Tuscaloosa, AL, pp. 508-573.
- Lupyan, G., and Clark, A. (2015). Words and the world: Predictive coding and the language-perception-cognition interface. *Current Directions in Psychological Science* 24, 279-284.
- Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature* 244, 522-523.
- Massaro, D., and Cohen, M. (1983). Phonological context in speech perception. *Perception & Psychophysics* 34, 338-348.
- McGurk, H., and MacDonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 746-748.
- Miller, G., and Nicely, P. (1955). An analysis of perceptual confusions among some English consonants. *The Journal of the Acoustical Society of America* 27, 338-352.

Nygaard, L. C., Sommers, M. S., and Pisoni, D. B. (1995). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception & Psychophysics* 57(7), 989-1001.

Pickett, J. M., and Pollack, I. (1963) Intelligibility of excerpts from fluent speech: Effects of rate of utterance and duration of excerpt. *Language and Speech* 6, 151-164.

Sekiyama, K., and Tohkura, Y. (1993). Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics* 21, 427-444.

Skipper, J. (2014). Echoes of the spoken past: How auditory cortex hears context during speech perception. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, 20130297.

Souza, P. E., Gehani, N., Wright, R. A., and McCloy, D. R. (2013). The advantage of knowing the talker. *Journal of the American Academy of Audiology* 24(8), 689-700.

Strand, E., and Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference*, Bielefeld, Germany. Mouton de Gruyter, Berlin, pp. 14-26.

Walden, B., Montgomery, A., Prosek, R., and Hawkins, D. (1990). Visual biasing of normal and impaired auditory speech perception. *Journal of Speech and Hearing Research* 33, 163-173.

Warren, R. (1970). Perceptual restoration of missing speech sounds. *Science* 167, 392-393

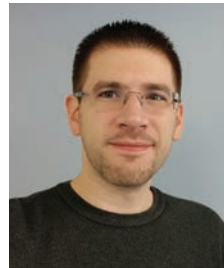
Weikum, W., Vouloumanos, A., Navarra, J., Soto-Foraco, S., Sebastian-Galies, N., and Werker, J. (2007). Visual language discrimination in infancy. *Science* 316, 1159.

Winn, M., Rhone, A., Chatterjee, M., and Idsardi, W. (2013). Auditory and visual context effects in phonetic perception by normal-hearing listen-

ers and listeners with cochlear implants. *Frontiers in Psychology: Auditory Cognitive Neuroscience* 4, 1-13.

Wright, R. A., and Souza, P. E. (2012). Comparing identification of standardized and regionally valid vowels. *Journal of Speech, Language, and Hearing Research* 55(1), 182-193.

## BioSketch



**Matthew B. Winn** is an assistant professor at the University of Washington, Seattle. His lab focuses on listening effort and speech perception in people with hearing loss as well as binaural hearing and basic auditory science. He received PhD and AuD (Doctor of Audiology) degrees from the University of Maryland, College Park, and worked as a clinical audiologist at the VA Medical Center in Washington, DC. His research is supported by the National Institute on Deafness and Other Communication Disorders. He is a member of the Acoustical Society of America and is on the Speech Communication and Psychological and Physiological Acoustics Committees. In the fall of 2018, he will join the faculty at the University of Minnesota, Minneapolis.

## NEWS from the Acoustical Society Foundation Fund

Although the Foundation Fund is technically one pool of funds, in reality, there are numerous discrete accounts set aside for different purposes. Four of these accounts support the named silver medals awarded by the Acoustical Society of America (ASA) for outstanding scientific contributions. You may find it interesting to know the history of these medal funds, which were established by designated generosity from donors and friends of the ASA.

**Pioneers of Underwater Acoustics Medal:** Originally funded by a \$500 contribution from the Westinghouse Electric Corporation in February 1959, the award was named in honor of these five pioneers in the field: H. J. W. Fay, R. A. Fessenden, H. C. Hayes, G. W. Pierce, and P. Langevin.

**von Békésy Medal:** Established in 1983 in recognition of “the generous gift of \$10,000 by the Belton Institute of Hearing Research to establish and endow a von Békésy Medal” in psychological and physiological acoustics.

**Trent-Crede Medal:** Established by the ASA in the mid-1960s to recognize contributions to the science of mechanical vibration and shock; it was funded from donor contributions and an initial fund drive.

**Sabine Medal:** Established with funds donated by the Acoustical Materials Association to recognize contributions in the field of architectural acoustics. The medal was first awarded in the fall of 1957.

Recipients are honored at the Plenary Sessions where the medals are awarded. The Foundation is very pleased to be able to support these commendations, and we are pursuing ways to increase the value of these and several other medals. A listing of the past recipients and details of the descriptions of the medals is available on the ASA website (see [acousticstoday.org/asa-awards](http://acousticstoday.org/asa-awards)).

Carl Rosenberg  
Chair, Acoustical Society Foundation Board

**Mission of the Acoustical Society Foundation Board:**

To support the mission of the ASA by developing financial resources for strategic initiatives and special purposes.

**ASFF** For more information, contact Carl Rosenberg at [crosenberg@acentech.com](mailto:crosenberg@acentech.com)