Hearing Research 381 (2019) 107773

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research Paper

Ecological cocktail party listening reveals the utility of extended high-frequency hearing

Brian B. Monson [a, b, *], Jenna Rock [a], Anneliese Schulz [a], Elissa Hoffman [a], Emily Buss [c]

[a] Department of Speech and Hearing Science, University of Illinois at Urbana-Champaign, United States
[b] Neuroscience Program, University of Illinois at Urbana-Champaign, United States
[c] Department of Otolaryngology/Head and Neck Surgery, University of North Carolina at Chapel Hill, United States

ARTICLE INFO

Article history:
Received 22 March 2019
Received in revised form
19 July 2019
Accepted 27 July 2019
Available online 3 August 2019

Keywords:
Cocktail party problem
Speech perception
High frequency hearing

ABSTRACT

A fundamental principle of neuroscience is that each species' and individual's sensory systems are tailored to meet the demands placed upon them by their environments and experiences. What has driven the upper limit of the human frequency range of hearing? The traditional view is that sensitivity to the highest frequencies (*i.e.*, beyond 8 kHz) facilitates localization of sounds in the environment. However, this has yet to be demonstrated for naturally occurring non-speech sounds. An alternative view is that, for social species such as humans, the biological relevance of conspecific vocalizations has driven the development and retention of auditory system features. Here, we provide evidence for the latter theory. We evaluated the contribution of extended high-frequency (EHF) hearing to common ecological speech perception tasks. We found that restricting access to EHFs reduced listeners' discrimination of talker head orientation by approximately 34%. Furthermore, access to EHFs significantly improved speech recognition under listening conditions in which the target talker's head was facing the listener while co-located background talkers faced away from the listener. Our findings raise the possibility that sensitivity to the highest audio frequencies fosters communication and socialization of the human species. These findings suggest that loss of sensitivity to the highest frequencies may lead to deficits in speech perception. Such EHF hearing loss typically goes undiagnosed, but is widespread among the middle-aged population.

1. Introduction

Different species have distinctive upper and lower limits to the frequency range of hearing (Fay, 1988; Heffner and Heffner, 2007; Heffner, 2004; Masterton et al., 1969), each presumably tailored to enable reproductive success according to the environmental demands for that species. For humans, the upper limit of the frequency range of hearing extends to approximately 20 kHz, with sensitivity to acoustic frequencies between 8 and 20 kHz being designated *extended* high-frequency (EHF) hearing.

What is the ecological utility of EHF hearing in humans? Although the EHF range is believed to be beneficial for some auditory tasks (*e.g.*, subjective judgments of sound and music quality; Monson et al., 2014a; Moore and Tan, 2003), the dominant view is that EHF hearing promotes survival and success by enhancing humans' ability to make elevation judgements and to resolve front/back discrepancies when localizing sound sources (Heffner and Heffner, 2008). Many studies have demonstrated that removing listeners' access to EHFs (*via* low-pass filtering) leads to increases in errors for elevation judgments and front/back distinctions (Best et al., 2005; Brungart and Simpson, 2009; Carlile et al., 1999; King and Oldfield, 1997). However, most demonstrations of this phenomenon have used synthetic stimuli (clicks and noise bursts) that exhibit artificially high levels of acoustical energy at high frequencies. Most natural sounds in the human environment exhibit energy roll off at higher frequencies, and it is not clear what ecological role EHF hearing has for localization of these sounds. Indeed, recent evidence indicates that front/back confusions abound when localizing more naturalistic sounds, even with access to EHFs (Derey et al., 2017).

It is widely believed that the EHF range plays little to no role in speech perception, being beyond the information-bearing traditional "speech bandwidth." One reason for this view is that foundational studies in speech and hearing science focused primarily on determining the frequency range of speech necessary and sufficient

Abbreviations: EHF, extended high frequency

* Corresponding author. 901 S Sixth St, Champaign, IL, 61820, United States.
E-mail address: monson@illinois.edu (B.B. Monson).

https://doi.org/10.1016/j.heares.2019.107773
0378-5955/

to reproduce highly intelligible speech for transmission over communication systems (Crandall and MacKenzie, 1922; Fletcher and Steinberg, 1930; Fletcher and Galt, 1950; Monson et al., 2014b). This work revealed that energy below 7 kHz was sufficient to achieve this objective, resulting in a general lack of study of the audibility and utility of speech spectral energy above this range. These studies did not often consider ecological conditions under which EHF information could potentially be useful. A related factor that may perpetuate this view of EHF and speech perception is the widely-held belief that little energy or acoustic structure exists in the speech signal beyond 8 kHz. As depicted in Fig. 1, EHF energy accounts for an appreciable portion of the speech spectrum (Levy et al., 2015; Monson et al., 2012b; Moore et al., 2008), suggesting potential utility for speech perception.

It has been proposed that, considering the social aptitude of humans and other species, the value of detecting and perceiving conspecific vocalizations exerts influence over the development and preservation of auditory system features (Manley, 2017; Theunissen and Elie, 2014). This suggests that, counter to the traditional view, the parameters of EHF hearing in humans may provide important insights about the mechanisms by which humans perceive speech. The present article addresses the biological significance of EHF hearing by considering the relationship between the human range of hearing, the frequency spectrum of speech, and the ecological demands placed on the auditory system. Accordingly, we examined the deficit produced by limiting access to EHF energy in speech for common ecological speech perception tasks, taking into consideration the fact that the emission of EHF energy in speech is highly dependent on the head orientation of the talker, relative to the listener (Chu and Warnock, 2002; Halkosaari et al., 2005; Kocon and Monson, 2018; Monson et al., 2012a; Rayleigh, 1908).

Whereas the ability of humans to localize sound sources, including speech, has been studied extensively (Middlebrooks, 2015), much less is understood of the human ability to detect the physical orientation of a sound source. In general, radiation patterns from a sound source, including speech radiating from the mouth, are frequency dependent, being more omnidirectional for lower frequency components and increasingly directional (*i.e.*, typically radiating toward the front of a talker) for higher frequency components (Chu and Warnock, 2002; Halkosaari et al., 2005; Kocon and Monson, 2018; Monson et al., 2012a; Rayleigh, 1908). This phenomenon makes the distribution of speech spectral energy at the ear of a listener dependent upon the physical orientation of a talker's head relative to the listener (Fig. 2A and B). Thus, just as visual information signals changes of a talker's head orientation (Wilson et al., 2000), acoustic cues are available to detect changes in a talker's head orientation. The potential for utilizing these acoustic cues for talker head orientation discrimination has been demonstrated with closed-set head-orientation identification tasks using a live talker (Edlund et al., 2012; Kato et al., 2010) or a rotating loudspeaker (Imbery et al., 2019). Furthermore, it has been proposed that auditory detection of head orientation is ecologically valuable for inferring when one is the intended recipient of an utterance or warning call (Neuhoff, 2003). Notably, because of the increasing directionality as frequency increases, EHFs are generally the most affected by changes in talker head orientation (see Fig. 2B), rendering EHF energy a potentially salient cue for discriminating head orientation. Thus we predicted that restricting access to EHFs would impair a listener's ability to detect changes in talker head orientation. We tested this prediction in Experiment 1.

One defining characteristic of the human auditory system is its ability to detect and recognize a target speech signal within a speech mixture, a task commonly referred to as the "cocktail party" problem (Cherry, 1953). This ability allows humans to successfully hold conversations in noisy environments like coffee shops, sporting events, and parties.

Traditional tests of speech-in-speech recognition use recordings obtained with a microphone located directly, or nearly directly, in front of the talkers (target talkers and interfering talkers). Presenting listeners with speech stimuli obtained in this way simulates a scenario in which both target and interfering talkers are directly facing the listener (Fig. 3A). This scenario results in masking for speech energy at both low and high frequencies. Under these conditions it has been demonstrated that EHF hearing (*i.e.*, access to speech spectral energy beyond 8 kHz) provides little to no benefit for target speech recognition in the presence of interfering speech (Levy et al., 2015; Moore et al., 2010). However, these listening conditions are not natural. A more natural scenario is that of a target talker facing the listener and interfering talkers facing in other directions (Fig. 3B). Given the directional nature of EHF energy emission in speech, this scenario results in preservation of low-frequency energy and attenuation of EHF energy associated with the interfering talkers. Based on this observation, we previously predicted that this listening condition would render EHF hearing more useful for target speech recognition (Monson et al., 2012a). We tested this prediction in Experiment 2.

We propose that maintaining the audibility of EHFs enables greater success in resolving the cocktail party problem, an ability critical to communication and socialization, and thus reproductive success and survival, of the species. In the present study, we find that EHF hearing holds greater utility for speech perception than is widely believed, lending support for the viewpoint that detection and perception of conspecific vocalizations has driven the upper limit of the human frequency range of hearing.

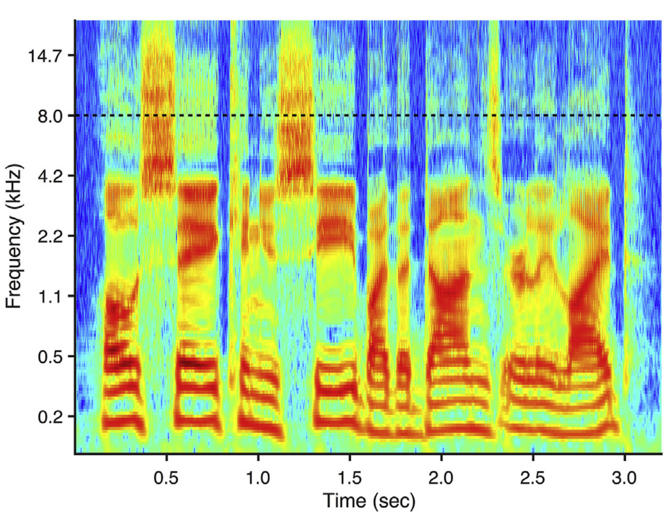

Fig. 1. Cochleagram of male speech. The phrase "Oh say, can you see by the dawn's early light" was spoken by a male talker. Appreciable energy and acoustic structure are apparent beyond 8 kHz (dotted line). To produce a perceptually relevant representation, data were plotted on a cochlear (equivalent rectangular bandwidth, ERB) frequency scale, using 1-ERB-wide filter bands with 50% overlap (Glasberg and Moore, 1990; McDermott and Simoncelli, 2011). (Transfer functions of the outer and middle ear were not incorporated.)

2. Experiment 1: Head orientation

2.1. Methods

Statistical analyses consisted of two-way repeated-measures analysis of variance (ANOVA), as described below. All statistical analyses were conducted using the *ezANOVA* function in R (R Core Team, 2018). Custom scripts written in MATLAB (MathWorks) were used for signal processing and experimental control. All recording materials and data for this study will be made available upon request. All experimental procedures were approved by the Institutional Review Board at the University of Illinoi at Urbana-Champaign.

2.1.1. Participants

Eighteen participants (two male, age 19–30 yr) participated in this experiment. Participants had normal hearing, as indicated by pure tone audiometric thresholds better than 20 dB HL in at least one ear for octave frequencies between 0.5 and 16 kHz and no history of hearing disorder.

2.1.2. Stimuli

The stimuli were recordings of the phrase “amend the slower page” uttered by two male and two female talkers, taken from a database of high-fidelity (44.1-kHz sampling rate, 16-bit precision) anechoic multi-channel recordings acquired with type I precision microphones surrounding the talker in 15° increments on a semicircle from 0° (directly in front) to 180° (directly behind) (Monson et al., 2012a). The recordings made across the microphone array capture the stimuli associated with a range of talker head orientations, from 0° (talker facing the listener) to 180° (talker facing away from the listener). Stimuli for the low-pass filtered condition were generated by low-pass filtering each recording with a 32-pole Butterworth filter with cutoff frequency of 8 kHz.

2.1.3. Procedure

Stimuli were presented to listeners seated in a sound-treated booth over a KRK Rokit 8 G3 loudspeaker at 70 dB SPL at 1 m directly in front of the listener. An adaptive one-up two-down, three-alternative forced-choice oddity task (Levitt, 1971) was used to measure thresholds for detecting the difference between a reference recording from 0° and a test recording from a different angle. The procedure was implemented using the AFC software (Ewert, 2013). Four tracks (one for each talker) were tested in separate runs. Two conditions were tested in separate blocks: full bandwidth and low-pass filtered at 8 kHz. Run order within each block was randomized, and block presentation order was randomized for each listener. Following a brief training block, each experimental run began with an easily detectable angle difference (135°). If 0° was reached and guessed correctly by chance, the trial was repeated until an error was made. The angle step size changed from 45° to 15° after the first two reversals. The angles at the last six reversals were averaged to obtain the detection thresholds. Feedback on accuracy was provided for the training block, but not for the experimental blocks.

2.2. Results

The average just noticeable differences (JNDs) in talker head orientation for the full bandwidth condition were 40° and 41° for female and male talkers, respectively (Fig. 2C). There was a main

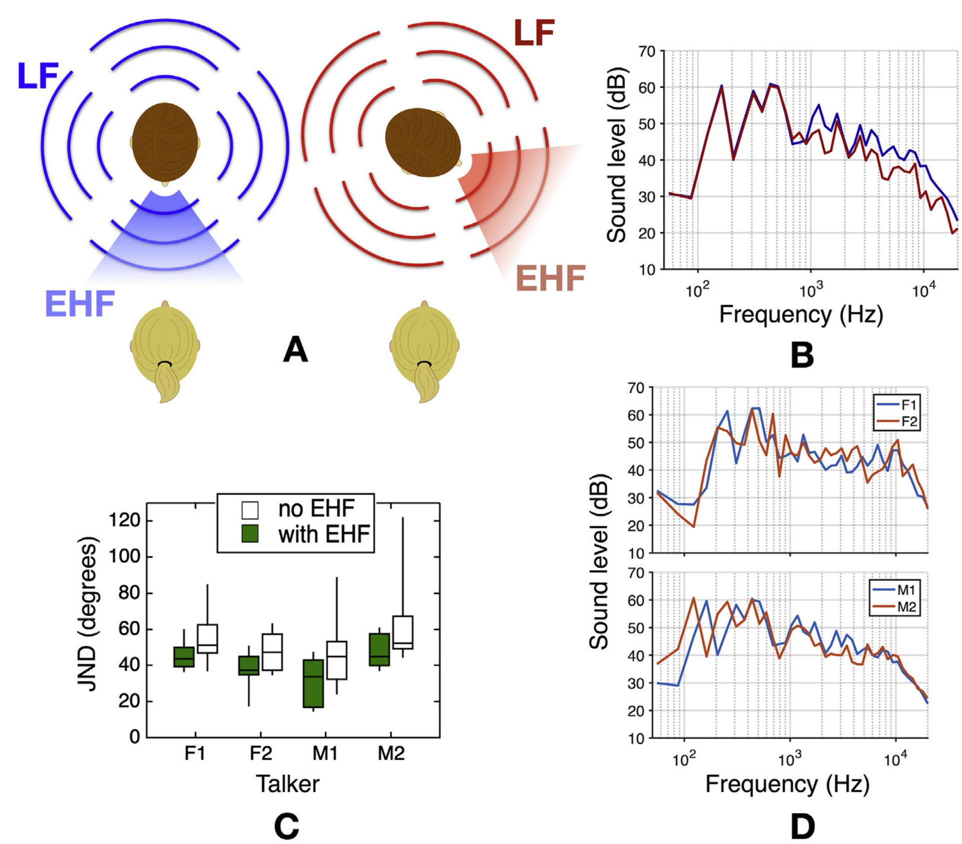

Fig. 2. A. Whereas low-frequency (LF) energy in speech radiates nearly omnidirectionally around a talker, EHF energy in speech radiates primarily toward the front of the talker. When the talker rotates their head from 0° (blue) to 60° (red) relative to the listener, LF energy shows little loss at the ear of the listener, whereas EHF energy exhibits appreciable loss. **B.** Speech spectra obtained from the location of the listener in panel A for talker M1 with head orientations of 0° (blue) and 60° (red) (Monson et al., 2012a). Data were plotted using 1-ERB-wide analysis bands. Energy losses of 6–10 dB are consistently observed at EHFs. **C.** Just noticeable differences (JND) in talker head orientation (relative to 0°) for speech uttered by two female and two male talkers. Average JNDs across all four talkers were 41° for full-bandwidth speech (green) and 55° for speech low-pass filtered at 8 kHz (white). **D.** Speech spectra (at 0°) for each of the four talkers, set to overall levels of 70 dB SPL. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

effect of filtering condition, with average JNDs of 56° and 53°, respectively, for male and female speech low-pass filtered at 8 kHz (two-way repeated measures ANOVA, $F(1,17) = 12.0$, $p = 0.003$). There was no main effect of talker sex ($F(1,17) = 0.2$, $p = 0.7$) and no interaction between filtering condition and talker sex ($F(1,17) = 0.9$, $p = 0.3$). Although there was variability across talkers, there was no clear relationship between the amount of EHF energy and the average JND for each talker (see Fig. 2D).

3. Experiment 2: Cocktail party listening

3.1. Methods

3.1.1. Participants

Twenty participants (five male, age 20–27 yr) participated in this experiment. Participants had normal hearing as indicated by pure tone audiometric thresholds better than 20 dB HL in at least one ear for octave frequencies between 0.5 Hz and 16 kHz and no history of hearing disorder.

3.1.2. Stimuli

The interferer stimulus was a two-female-talker babble created using 45° and 60° recordings from the same multichannel multi-angle database as used for Experiment 1 (Monson et al., 2012a). A semantically unpredictable speech signal of approximately 2 min' duration was generated for each talker by extracting individual words from 20 phrases uttered in the original recording and concatenating them in a randomized order. The unpredictable speech signals for each talker were then summed to create the babble stimulus. Target speech stimuli were the BKB sentences (Bench et al., 1979) uttered by a single female talker, recorded in a sound-treated booth using a type I precision microphone located at 0°, with 44.1-kHz sampling rate and 16-bit precision. For the low-pass filtered condition, all stimuli were low-pass filtered using a 32-pole Butterworth filter with cutoff frequency of 8 kHz.

3.1.3. Procedure

Stimuli were presented to listeners seated in a sound-treated booth over a KRK Rokit 8 G3 loudspeaker at 1 m directly in front of the listener. The level of the two-talker interferer was set at 70 dB SPL at 1 m, while the level of the target signal (determining the signal-to-noise ratio, SNR) was adaptively varied. Two interleaved adaptive tracks were utilized, both using a one-down, one-up adaptive rule. For one track, the SNR was reduced if the listener got one or more words correct; otherwise the SNR was increased. For the other track, the SNR was reduced if the listener got all words or all but one word correct; otherwise the SNR was increased. Both adaptive tracks started at 4 dB SNR. The SNR was initially adjusted in steps of 4 dB, and then by 2 dB after the first reversal. Each of the two tracks comprised 32 sentences. Word level data from the two tracks were combined and fitted with a logit function with asymptotes at 0 and 100% correct. One advantage to this approach is that it provides an estimate of both the psychometric function slope and the speech reception threshold (SRT, defined as the SNR associated with 50% correct), characterizing performance across a range of SNRs. Data fits were associated with r^2 values ranging from 0.53 to 0.99, with a median value of 0.87.

Two filtering conditions were tested: full band vs. all stimuli low-pass filtered at 8 kHz. Two interferer head angle conditions were tested: both interferers facing 45° or both interferers facing 60° relative to the target talker (see Fig. 3B and D). After a brief training block consisting of 16 sentences, the four conditions (2 filtering conditions × 2 interferer head angles) were tested in separate blocks with block order randomized across participants. The starting sentence list number was randomized for each participant and continued in numerical order of the BKB sentence lists.

3.2. Results

There was a main effect of low-pass filtering, with reductions in the mean SRT of 1.4 dB (1.6 dB, median) and 2 dB (2.5 dB, median) for the 45° and 60° conditions, respectively (two-way repeated measures ANOVA, $F(1,19) = 20.8$, $p < 0.001$; Fig. 3C). There was a main effect of interferer head orientation angle, with better performance for the 60° condition ($F(1,19) = 17.5$, $p < 0.001$), and no interaction between filtering condition and interferer head orientation ($F(1,19) = 0.67$, $p = 0.4$). There was no difference in psychometric slopes across filtering conditions ($F(1,19) = 0.65$, $p = 0.4$) or angle conditions ($F(1,19) = 0.05$, $p = 0.8$).

4. General discussion

A particular species' and individual's auditory brain develops sensitivity to the sounds that matter most for their survival (*e.g.*, the vocalizations of predators, prey, and mates (Hauser, 1996; Hoy, 1992; Manley, 2017; Theunissen and Elie, 2014); and success (*e.g.*, the spectral detail of phonemes of one's native language; (Kuhl et al., 1992; Liberman et al., 1957; Werker and Tees, 1984). The observation that humans display sensitivity to speech spectral energy at EHFs raises the possibility that this energy provides information regarding the speech signal. Our results reveal multiple uses of EHF energy for speech perception.

4.1. Head orientation

On average, listeners were to be sensitive to changes of approximately 41° in talker head orientation using only auditory cues for the full-bandwidth stimuli. Not surprisingly, this sensitivity is much poorer than that for visual cues for changes in head orientation, for which humans can discriminate changes of only a few degrees (Wilson et al., 2000). However, the ability to detect a change in talker head orientation based solely on auditory cues can be useful when visual cues are not available, for example, in darkness or when the talker is not in a listener's field of view.

Auditory detection of head orientation can be valuable for inferring when one is the intended recipient of an utterance or warning call (Neuhoff, 2003). Although a JND of 41° seems large, it may be that this position represents the boundary at which a talker can be judged to be facing the listener (*i.e.*, speaking to the listener) or facing away from the listener (*i.e.*, speaking to someone else). When access to EHFs was restricted, there was a reduction in discrimination performance by approximately 14° (34%). Our finding that head orientation discrimination is better with EHF hearing suggests that EHF sensitivity could also improve one's ability to determine when one is the intended recipient of a vocalization, which is of communicative value.

4.2. Cocktail party listening

To our knowledge, only one other study has directly assessed the effects of interfering talkers' head orientation on speech-in-speech recognition (Strelcyk et al., 2014; however, see also Moore and Popelka, 2013; Plomp and Mimpen, 1981). Strelcyk et al. (2014) reported an improvement in target speech recognition when the interfering talkers' heads were rotated from 0° to 105° relative to the listener. However, the stimuli used for that study were all low-pass filtered at 8 kHz, precluding any inferences regarding the role of EHF hearing. Using the traditional experimental arrangement (*i.e.*, with all talkers facing the listener; see Fig. 3A), others have

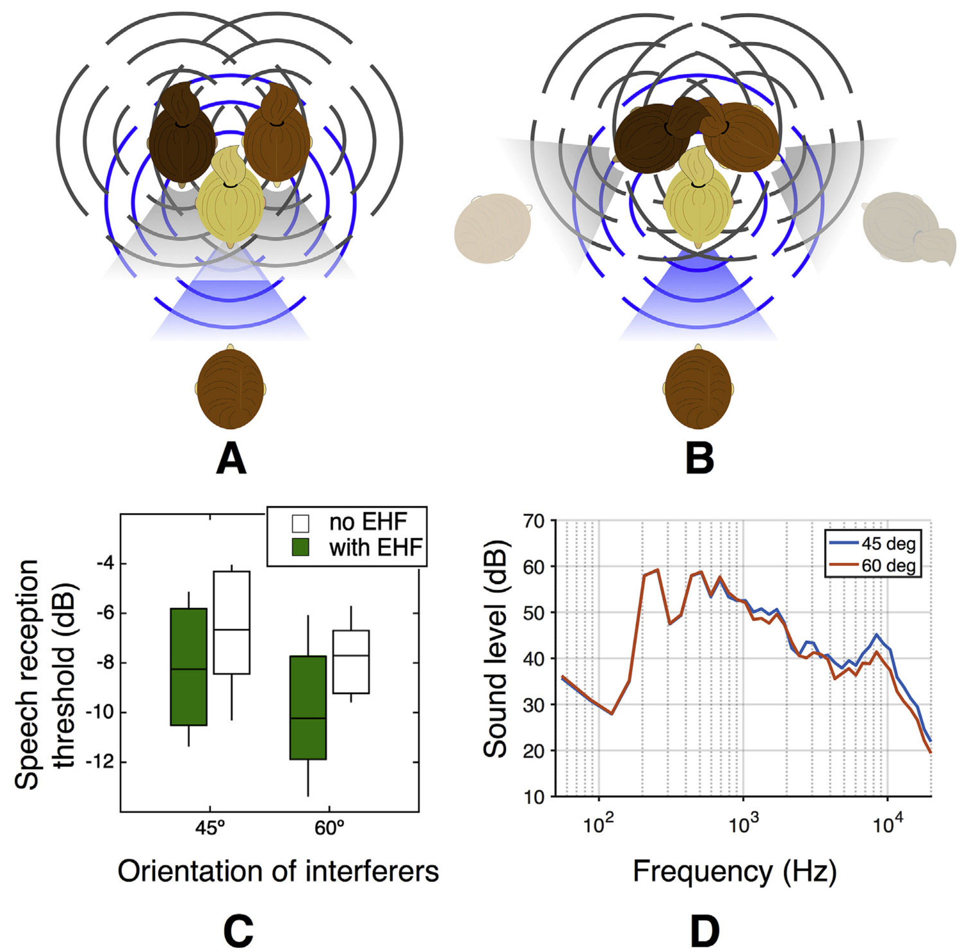

Fig. 3. A. Traditionally, speech-in-speech recognition is assessed by simulating a scenario in which both the target (blue) and interferers (gray) are facing the listener. **B.** The more ecological scenario used in the present study is that of the target talker facing the listener and rotated interfering talkers (*i.e.*, talking to other [faded] hypothetical listeners). Interferers are depicted with 60° head rotations. For easier visibility, the co-located interferers have been given slight separation in panels A and B. **C.** Speech reception thresholds for the listening scenario depicted in panel B. **D.** Spectra comparing 45° and 60° two-talker babble. Data were plotted using 1-ERB-wide analysis bands and set to overall levels of 70 dB SPL. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

evaluated the effect of low-pass filtering at 7.5 kHz (Moore et al., 2010) and 8 kHz (Levy et al., 2015) on speech-in-speech recognition, but little to no benefit was observed for the full-bandwidth condition (which was limited to 10 kHz). In particular, no difference was observed for speech recognition between speech low-pass filtered at 7.5 kHz and 10 kHz, whether target and interfering speech were co-located or spatially separated (Moore et al., 2010). Using spatially separated target and interferers (but again with all talkers facing the listener), no difference in speech recognition was observed between low-pass filtering conditions of 8 kHz and 10 kHz, nor between 6 kHz and 8 kHz, although a 1.3-dB improvement was observed when comparing 6 kHz and 10 kHz low-pass filtering conditions (Levy et al., 2015). This result suggests that the EHF band between 8 and 10 kHz may have some utility when combined with the 6–8 kHz band, if interfering talkers are spatially separated from the target talker, and if all talkers are facing the listener.

In the present study, more ecologically relevant aspects were incorporated in the experimental setup to assess whether an effect of EHF information could be measured. Specifically, we found that when co-located interfering talkers were rotated away from the listener, access to EHFs significantly improved speech recognition. The improvements observed under these conditions were equal to or greater than that reported previously for the 6–10 kHz band with spatially separated interferers in the traditional experimental arrangement (Levy et al., 2015). Our findings provide some explanation for previous observations that individuals who suffer from poor speech recognition in noisy environments also tend to have EHF hearing loss (Badri et al., 2011). Together, these findings are striking because they demonstrate a deficit associated with EHF hearing loss that would occur under natural conditions. Age-related hearing loss at EHFs is a very common condition known to begin in young adulthood (Green et al., 1987; Stelmachowicz et al., 1989). Critically, however, this condition typically goes undiagnosed because EHF sensitivity testing is not part of routine audiological clinical assessments at present (Moore et al., 2017). The inclusion of EHF assessments in routine exams may provide valuable information to identify listeners at risk of experiencing listening difficulties in noisy environments. It is noteworthy that EHF hearing loss has been proposed as a potential marker for cochlear synaptopathy or “hidden hearing loss” in individuals with otherwise normal clinical audiograms (Prendergast et al., 2017). Our results raise the possibility, however, that EHF hearing loss *per se* might give rise to listening difficulties that could potentially be erroneously interpreted as “hidden hearing loss” or some other auditory disorder.

Our results have clear implications for hearing aids, cell phones, and other communication devices that typically do not transmit sounds in the EHF range. For example, most current hearing aids do not provide amplification for most of the EHF range, and hearing-aid users consistently complain about poor performance of hearing aids for speech recognition in noisy environments (Kochkin, 2010). Our results suggest that providing effective amplification and transmission beyond 8 kHz could be beneficial for speech recognition, particularly in the presence of competing speech signals. Hearing aids that do provide amplification for at least some of the EHF range (up to 10 kHz) show improvements in users' subjective ratings of hearing-aid performance (Arbogast et al., 2019).

While we have demonstrated the utility of EHF hearing in cocktail party listening, some questions remain. Does EHF speech spectral energy provide phonetic information *per se* under these listening conditions, or does EHF energy simply serve as a grouping cue that increases the utility of low-frequency information? The temporal coherence of different features of a common sound enables those features to be grouped together into a single auditory "object" or "stream", thereby improving auditory scene analysis (Shamma et al., 2011). Because EHF speech energy is at least partly coherent with low-frequency speech energy (see Fig. 1; (Crouzet and Ainsworth, 2001), it may be that access to EHF energy from the target speech facilitates segregation of low-frequency phonetic information from the interfering speech. On the other hand, because EHF speech spectral energy *per se* does provide some phonetic information (Berlin et al., 1978; Lippmann, 1996; Vitela et al., 2015), it is possible that this additional phonetic information contributed to the benefit we observed. Furthermore, although we incorporated some ecologically relevant aspects of cocktail party listening in the present study, it will be important to assess other ecological arrangements (*e.g.*, increasing the number of interferers, spatially separated rotated interferers at differing distances or elevations, interferers behind the head; Martin et al., 2012).

4.3. Conclusions

The observation that EHF hearing is beneficial to head orientation discrimination and speech recognition suggests that its preservation fosters communication and socialization of the human species. In particular, EHF cues accessible under more ecologically relevant conditions led to improvements in cocktail party listening. More broadly, in conjunction with prior work revealing that ample EHF spectral energy is produced during speech and that it displays phonetically distinct spectral features (Monson et al., 2012b; Vitela et al., 2015), the present findings provide support for the idea that basic features of the human auditory nervous system and vocal mechanism are tuned to each other. There could be other important benefits of EHF hearing in complex listening environments, such as sound source localization. Although this has yet to be widely demonstrated for natural sounds, one notable exception demonstrated the utility of EHF cues for improved front/back discrimination of human speech (Best et al., 2005), an observation also made by Lord Rayleigh over a century ago (Rayleigh, 1908). These past findings bolster the argument that EHF sensitivity is particularly useful for perception of speech. Other potential benefits of EHF hearing that warrant further exploration include accelerated word learning for children (Pittman, 2008; Stelmachowicz et al., 2007), who possess superior sensitivity in the EHF range (Stelmachowicz et al., 1989), and perception of speech, voice, and music quality, which is known to be affected by EHF (*i.e.*, treble) audibility (Monson et al., 2014a; Moore and Tan, 2003).

Author contributions

B.B.M. designed Experiment 1. B.B.M. and E.B. designed Experiment 2 and analyzed data. J.R., A.S., and E.H. collected data. B.B.M. and J.R. drafted the manuscript. All authors reviewed and edited the manuscript.

Acknowledgements

The authors thank David Frazier, Olivia Godnik, Melanie Flores, Elana Hunt, Joanna Way, and Molly Cull for assistance with data collection. Kellie Halloran assisted with figure design. This work was supported by NIH Grant R01DC000397.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heares.2019.107773.

References

Arbogast, T.L., Moore, B.C.J., Puria, S., Dundas, D., Brimacombe, J., Edwards, B., Carr Levy, S., 2019. Achieved gain and subjective outcomes for a wide-bandwidth contact hearing aid fitted using CAM2. Ear Hear. 40, 741–756.
Badri, R., Siegel, J.H., Wright, B.A., 2011. Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. J. Acoust. Soc. Am. 129, 852–863.
Bench, J., Kowal, A., Bamford, J., 1979. The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. Br. J. Audiol. 13, 108–112.
Berlin, C.I., Wexler, K.F., Jerger, J.F., Halperin, H.R., Smith, S., 1978. Superior ultra-audiometric hearing: a new type of hearing loss which correlates highly with unusually good speech in the "profoundly deaf. Otolaryngology 86. ORL-111-116.
Best, V., Carlile, S., Jin, C., van Schaik, A., 2005. The role of high frequencies in speech localization. J. Acoust. Soc. Am. 118, 353–363.
Brungart, D.S., Simpson, B.D., 2009. Effects of bandwidth on auditory localization with a noise masker. J. Acoust. Soc. Am. 126, 3199–3208.
Carlile, S., Delaney, S., Corderoy, A., 1999. The localisation of spectrally restricted sounds by human listeners. Hear. Res. 128, 175–189.
Cherry, E.C., 1953. Some experiments on the recognition of speech, with one and with two ears. J. Acoust. Soc. Am. 25, 975–979.
Chu, W., Warnock, A., 2002. Detailed Directivity of Sound Fields Around Human Talkers. Institute for Research in Construction, Ottawa ON, Canada.
Crandall, I.B., MacKenzie, D., 1922. Analysis of the energy distribution in speech. Phys. Rev. 19, 221–232.
Crouzet, O., Ainsworth, W.A., 2001. On the various influences of envelope information on the perception of speech in adverse conditions: an analysis of between-channel envelope correlation. In: Workshop on Consistent and Reliable Acoustic Cues for Sound Analysis. Aalborg, Denmark.
Derey, K., Rauschecker, J.P., Formisano, E., Valente, G., de Gelder, B., 2017. Localization of complex sounds is modulated by behavioral relevance and sound category. J. Acoust. Soc. Am. 142, 1757–1773.
Edlund, J., Heldner, M., Gustafson, J., 2012. On the effect of the acoustic environment on the accuracy of perception of speaker orientation from auditory cues alone. In: 13th Annual Conference of the International Speech Communication Association 2012, INTERSPEECH 2012. Curran Associates, Inc., pp. 1482–1485
Ewert, S.D., 2013. AFC—a modular framework for running psychoacoustic experiments and computational perception models. In: Proceedings of the International Conference on Acoustics AIA-DAGA, pp. 1326–1329.
Fay, R.R., 1988. Comparative psychoacoustics. Hear. Res. 34, 295–305.
Fletcher, H., Steinberg, J.C., 1930. Articulation testing methods. J. Acoust. Soc. Am. 1, A1–A48.
Fletcher, H., Galt, R.H., 1950. The perception of speech and its relation to telephony. J. Acoust. Soc. Am. 22, 89–151.
Glasberg, B.R., Moore, B.C.J., 1990. Derivation of auditory filter shapes from notched-noise data. Hear. Res. 47, 103–138.
Green, D.M., Kidd Jr., G., Stevens, K.N., 1987. High-frequency audiometric assessment of a young adult population. J. Acoust. Soc. Am. 81, 485–494.
Halkosaari, T., Vaalgamaa, M., Karjalainen, M., 2005. Directivity of artificial and human speech. J. Audio Eng. Soc. 53, 620–631.
Hauser, M.D., 1996. The Evolution of Communication. MIT press, Cambridge, MA.
Heffner, H.E., Heffner, R.S., 2007. Hearing ranges of laboratory animals. J. Am. Assoc. Lab. Anim. Sci. 46, 20–22.
Heffner, H.E., Heffner, R.S., 2008. High-frequency hearing. In: Dallos, P., Oertel, D., Hoy, R. (Eds.), Handbook of the Senses: Audition. Elsevier, New York, pp. 55–60.
Heffner, R.S., 2004. Primate hearing from a mammalian perspective. Anat. Rec. A. Discov. Mol. Cell. Evol. Biol. 281, 1111–1122.
Hoy, R.R., 1992. The Evolution of Hearing in Insects as an Adaptation to Predation from Bats, the Evolutionary Biology of Hearing. Springer, New York, NY, pp. 115–129.
Imbery, C., Franz, S., van de Par, S., Bitzer, J., 2019. Auditory facing angle perception: the effect of different source positions in a real and an anechoic environment.

Acta Acust United Ac 105, 492–505.
Kato, H., Takemoto, H., Nishimura, R., Mokhtari, P., 2010. Spatial acoustic cues for the auditory perception of speaker's facing direction. In: Proc. of 20th International Congress on Acoustics. ICA, 2010.
King, R.B., Oldfield, S.R., 1997. The impact of signal bandwidth on auditory localization: implications for the design of three-dimensional audio displays. Hum. Factors 39, 287–295.
Kochkin, S., 2010. MarkeTrak VIII: consumer satisfaction with hearing aids is slowly increasing. Hear. J. 63, 19–20.
Kocon, P., Monson, B.B., 2018. Horizontal directivity patterns differ between vowels extracted from running speech. J. Acoust. Soc. Am. 144, EL7.
Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. Science 255, 606–608.
Levitt, H., 1971. Transformed up-down methods in psychoacoustics. J. Acoust. Soc. Am. 49, 467–477.
Levy, S.C., Freed, D.J., Nilsson, M., Moore, B.C.J., Puria, S., 2015. Extended high-frequency bandwidth improves speech reception in the presence of spatially separated masking speech. Ear Hear. 36, e214–e224.
Liberman, A.M., Harris, K.S., Hoffman, H.S., Griffith, B.C., 1957. The discrimination of speech sounds within and across phoneme boundaries. J. Exp. Psychol. 54, 358–368.
Lippmann, R.P., 1996. Accurate consonant perception without mid-frequency speech energy. IEEE Trans. Speech Audio Process. 4, 66–69.
Manley, G.A., 2017. Comparative auditory neuroscience: understanding the evolution and function of ears. J. Assoc. Res. Otolaryngol. 18, 1–24.
Martin, R.L., McAnally, K.I., Bolia, R.S., Eberle, G., Brungart, D.S., 2012. Spatial release from speech-on-speech masking in the median sagittal plane. J. Acoust. Soc. Am. 131, 378–385.
Masterton, B., Heffner, H., Ravizza, R., 1969. The evolution of human hearing. J. Acoust. Soc. Am. 45, 966–985.
McDermott, J.H., Simoncelli, E.P., 2011. Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. Neuron 71, 926–940.
Middlebrooks, J.C., 2015. Sound localization. Handb. Clin. Neurol. 129, 99–116.
Monson, B.B., Hunter, E.J., Story, B.H., 2012a. Horizontal directivity of low- and high-frequency energy in speech and singing. J. Acoust. Soc. Am. 132, 433–441.
Monson, B.B., Lotto, A.J., Story, B.H., 2012b. Analysis of high-frequency energy in long-term average spectra of singing, speech, and voiceless fricatives. J. Acoust. Soc. Am. 132, 1754–1764.
Monson, B.B., Lotto, A.J., Story, B.H., 2014a. Detection of high-frequency energy level changes in speech and singing. J. Acoust. Soc. Am. 135, 400–406.
Monson, B.B., Hunter, E.J., Lotto, A.J., Story, B.H., 2014b. The perceptual significance of high-frequency energy in the human voice. Front. Psychol. 5, 587.
Moore, B.C.J., Tan, C.T., 2003. Perceived naturalness of spectrally distorted speech and music. J. Acoust. Soc. Am. 114, 408–419.
Moore, B.C.J., Popelka, G.R., 2013. Preliminary comparison of bone-anchored hearing instruments and a dental device as treatments for unilateral hearing loss. Int. J. Audiol. 52, 678–686.
Moore, B.C.J., Fullgrabe, C., Stone, M.A., 2010. Effect of spatial separation, extended bandwidth, and compression speed on intelligibility in a competing-speech task. J. Acoust. Soc. Am. 128, 360–371.
Moore, B.C.J., Stone, M.A., Fullgrabe, C., Glasberg, B.R., Puria, S., 2008. Spectro-temporal characteristics of speech at high frequencies, and the potential for restoration of audibility to people with mild-to-moderate hearing loss. Ear Hear. 29, 907–922.
Moore, D., Hunter, L., Munro, K., 2017. Benefits of extended high-frequency audiometry for everyone. Hear. J. 70, 50–52.
Neuhoff, J.G., 2003. Twist and shout: audible facing angles and dynamic rotation. Ecol. Psychol. 15, 335–351.
Pittman, A.L., 2008. Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. J. Speech Lang. Hear. Res. 51, 785–797.
Plomp, R., Mimpen, A.M., 1981. Effect of the orientation of the speakers head and the azimuth of a noise source on the speech-reception threshold for sentences. Acustica 48, 325–328.
Prendergast, G., Millman, R.E., Guest, H., Munro, K.J., Kluk, K., Dewey, R.S., Hall, D.A., Heinz, M.G., Plack, C.J., 2017. Effects of noise exposure on young adults with normal audiograms II: behavioral measures. Hear. Res. 356, 74–86.
R Core Team, 2018. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rayleigh, 1908. Acoustical notes. - VIII. Philos. Mag. 16, 235–246.
Shamma, S.A., Elhilali, M., Micheyl, C., 2011. Temporal coherence and attention in auditory scene analysis. Trends Neurosci. 34, 114–123.
Stelmachowicz, P.G., Lewis, D.E., Choi, S., Hoover, B., 2007. Effect of stimulus bandwidth on auditory skills in normal-hearing and hearing-impaired children. Ear Hear. 28, 483–494.
Stelmachowicz, P.G., Beauchaine, K.A., Kalberer, A., Kelly, W.J., Jesteadt, W., 1989. High-frequency audiometry: test reliability and procedural considerations. J. Acoust. Soc. Am. 85, 879–887.
Strelcyk, O., Pentony, S., Kalluri, S., Edwards, B., 2014. Effects of interferer facing orientation on speech perception by normal-hearing and hearing-impaired listeners. J. Acoust. Soc. Am. 135, 1419–1432.
Theunissen, F.E., Elie, J.E., 2014. Neural processing of natural sounds. Nat. Rev. Neurosci. 15, 355–366.
Vitela, A.D., Monson, B.B., Lotto, A.J., 2015. Phoneme categorization relying solely on high-frequency energy. J. Acoust. Soc. Am. 137, EL65–EL70.
Werker, J.F., Tees, R.C., 1984. Cross-language speech-perception - evidence for perceptual reorganization during the 1st year of life. Infant Behav. Dev. 7, 49–63.
Wilson, H.R., Wilkinson, F., Lin, L.M., Castillo, M., 2000. Perception of head orientation. Vis. Res. 40, 459–472.