Adaptive spatial filtering improves speech reception in noise while preserving binaural cues

Susan R. S. Bissmeyer[a] and Raymond L. Goldsworthy
Caruso Department of Otolaryngology, Caruso Center for Childhood Communication, University of Southern California, 806 West Adams Boulevard, Los Angeles, California 90007, USA

(Received 25 November 2015; revised 8 August 2017; accepted 28 August 2017; published online 19 September 2017)

Hearing loss greatly reduces an individual's ability to comprehend speech in the presence of background noise. Over the past decades, numerous signal-processing algorithms have been developed to improve speech reception in these situations for cochlear implant and hearing aid users. One challenge is to reduce background noise while not introducing interaural distortion that would degrade binaural hearing. The present study evaluates a noise reduction algorithm, referred to as binaural Fennec, that was designed to improve speech reception in background noise while preserving binaural cues. Speech reception thresholds were measured for normal-hearing listeners in a simulated environment with target speech generated in front of the listener and background noise originating 90° to the right of the listener. Lateralization thresholds were also measured in the presence of background noise. These measures were conducted in anechoic and reverberant environments. Results indicate that the algorithm improved speech reception thresholds, even in highly reverberant environments. Results indicate that the algorithm also improved lateralization thresholds for the anechoic environment while not affecting lateralization thresholds for the reverberant environments. These results provide clear evidence that this algorithm can improve speech reception in background noise while preserving binaural cues used to lateralize sound.
 [http://dx.doi.org/10.1121/1.5002691]

[VMR] Pages: 1441–1453

I. INTRODUCTION

People with hearing loss struggle to take advantage of situations where background noise is spatially separate from desired speech (e.g., Arbogast *et al.,* 2005; Marrone *et al.,* 2008b,c; Neher *et al.,* 2011; Best *et al.,* 2012; Woods *et al.,* 2013; Kidd *et al.,* 2015). Studies have compared speech reception when the desired speech and unwanted masker have the same location versus when they are spatially separate (Freyman *et al.,* 1999; Brungart, 2001; Freyman *et al.,* 2001; Hawley *et al.,* 2004). These studies have demonstrated that people with normal hearing have better speech reception when the masker is spatially separate from the desired speech, an effect referred to as "spatial release from masking" (Hirsh, 1950; Hawley *et al.,* 2004; Kidd *et al.,* 2010; Swaminathan *et al.,* 2016). People with hearing loss, even those treated with cochlear implants or hearing aids, may exhibit little to no spatial release from masking (Marrone *et al.,* 2008b; Loizou *et al.,* 2009; Rothpletz *et al.,* 2012). A long-term clinical goal is to restore sufficient hearing to those with hearing loss so that they may benefit from such spatial release from masking. However, a more immediate solution to this problem is the use of multiple microphone signal processing, or directional microphones, to provide spatial filtering of a sound before it is presented to the hearing-impaired listener.

There is a substantial history concerning multiple microphone signal processing for selectively enhancing sounds based on spatial location (for reviews: Van Veen and Buckley, 1988; Brandstein and Ward, 2001). Multiple microphone signal processing can be subdivided into fixed and adaptive algorithms. In fixed algorithms, linear combinations of microphones, or microphone ports, are used to form a spatial filter that is independent of the input acoustics. Common examples of fixed spatial filters include cardioid and dipole response patterns, which have been implemented and shown to provide consistent speech reception benefits for both cochlear implant and hearing aid users (Soede *et al.,* 1993a; Soede *et al.,* 1993b; Stadler and Rabinowitz, 1993; Kates, 1993; Desloge *et al.,* 1997; Chung, 2004; Chung *et al.,* 2004; Chung *et al.,* 2006; Chung and Zeng, 2009).

Closely related to fixed algorithms, the earliest adaptive multiple microphone algorithms were developed to dynamically adjust filter weights used for combining the microphone signals to minimize output noise power. These early adaptive beamforming algorithms have generally been referred to as null-steering beamformers (Frost, 1972; Griffiths and Jim, 1982), which have evolved over the years and have been used on many clinical cochlear implant and hearing aid processors (Greenberg and Zurek, 1992; Kompis and Dillier, 1994; Kates and Weiss, 1996; Welker *et al.,* 1997; Berghe and Wouters, 1998; Spriet *et al.,* 2007; Hersbach *et al.,* 2012; Kokkinakis *et al.,* 2012). Such null-steering beamformers have been shown to provide speech reception benefits in background noise; but benefits quickly diminish with increasing reverberation and/or number of noise sources (Greenberg and Zurek, 1992; van Hoesel and Clark, 1995; Hamacher *et al.,* 1997; Wouters and Vanden

[a]Electronic mail: ssubrahm@usc.edu

Berghe, 2001; Kokkinakis and Loizou, 2010; Hazrati and Loizou, 2012; Desmond *et al.,* 2014).

In contrast to null-steering beamformers, a relatively new class of beamforming algorithms has been developed using a fundamentally different approach. Instead of slowly adapting the steering of spatial nulls, these algorithms use relatively rapid spectrotemporal signal analysis to determine which components are dominated by target or by masker energy and then preserve or attenuate the components accordingly. This general approach was inspired by models of binaural hearing (Jeffress, 1948) leading to the pioneering work of Kollmeier and colleagues (Kollmeier *et al.,* 1993; Kollmeier and Koch, 1994). Since the signal processing objective of this class of beamformer is to isolate the target speech and to suppress all other sounds, it is appropriately referred to as target-isolating beamformers. Lockwood *et al.* (2004) conducted a systematic study that included a fixed beamformer, two null-steering beamformers, and two target-isolating beamformers and demonstrated that the target-isolating beamformers were relatively robust to reverberation compared to null-steering beamformers. Further developing this class of beamformer, Goldsworthy and colleagues (Goldsworthy *et al.,* 2014; Goldsworthy, 2014) developed an algorithm, referred to as Fennec, based on analysis of inter-microphone phase differences between closely spaced microphones in a behind-the-ear capsule. They demonstrated that the Fennec algorithm provided speech reception benefits for cochlear implant users even with moderate levels of reverberation and as many as 11 noise sources.

The purpose of the study presented in this article was to evaluate a binaural version of the Fennec algorithm designed to improve the target to masker ratio when the target is in front of the listener and the masker is spatially separate, while preserving binaural cues. Previous studies that have considered configuring spatial filtering to preserve binaural cues have had varying levels of success. Desloge *et al.* (1997) demonstrated that fixed beamforming provides a degree of noise suppression while preserving binaural cues. However, subsequent work concerning adaptive spatial filtering has generally been unsuccessful towards achieving both noise suppression and preservation of binaural cues. One modification put forth by Welker *et al.* (1997) and revisited by Kidd *et al.* (2015) was to divide the acoustic spectrum into two regions and to implement noise suppression in one spectral region while preserving binaural cues in the other spectral region. That approach circumvents the problem to an extent by either performing spatial filtering within a spectral region or preserving the binaural cues, but both objectives are not achieved for any spectral region.

A different approach for spatial filtering based on binaurally situated microphones using adaptive Wiener filtering was evaluated for noise suppression while preserving binaural cues (Klasen *et al.,* 2006; Klasen *et al.,* 2007; Bogaert *et al.,* 2007; Bogaert *et al.,* 2008; Szurley *et al.,* 2016). This approach, rather than dividing the acoustic spectrum into processed and unprocessed regions, introduced a cost function to control the relative degree of noise suppression to preserve binaural cues (Klasen *et al.,* 2006). This approach had limited success, the limiting factor being that to preserve binaural cues a substantial loss in noise suppression had to occur (Bogaert *et al.,* 2007). A second limitation of this approach was that while the modification preserves the binaural cues associated with the target sound, it does not preserve the binaural cues associated with the environmental noise sources (Klasen *et al.,* 2007; Bogaert *et al.,* 2008; Kokkinakis *et al.,* 2012). A more recent study of this approach investigated the use of a remote microphone with a high target to masker ratio to control the adaptive procedure to achieve both noise suppression and binaural cue preservation (Szurley *et al.,* 2016). While a theoretically important step, it is not practical in cochlear implant and hearing aid applications to presume that the listener will have available a remote microphone with a clean representation of the target sound.

More recently, an approach using a signal-to-noise ratio estimator that controls a binary decision mask was evaluated for both providing noise suppression while preserving binaural cues (Thiemann *et al.,* 2016). That evaluation demonstrated that a constrained target-isolating beamformer could be successfully configured to provide noise suppression while preserving binaural cues sufficient to convey sound locations for both the target speech as well as environmental noise sources.

Inherent ear asymmetries, such as auditory nerve survival or physical differences, and asymmetries introduced by the clinical processors, such as adjusted gain control, can confound interaural level and timing cues for the hearing-impaired listener. Since these are a significant clinical and signal processing problem, it is essential that the spatial beamforming algorithm does not introduce any additional interaural distortions. This goal can be achieved by requiring that the spatial beamforming algorithm applies identical spectrotemporal noise attenuation to the left and right ear devices at any given moment for any given frequency component. In that manner, if the left and right ear devices at some moment in time and at some frequency, have specific interaural level and timing differences, then attenuating the left and right ears by the same amount will preserve the original interaural differences.

There are numerous ways in which independent left and right ear beamformers could be combined to produce a coordinated output. The approach considered in this article is straightforward with the attenuation functions of left and right ear beamformers averaged to produce a single attenuation function that is jointly applied to left- and right-ear microphone signals. Other approaches based on acoustic analysis and dynamic switching to the ear with better signal-to-noise ratio might be developed to enhance this approach. In the case of dynamic switching, the algorithm could determine which ear has the higher SNR and give more weight to that ear in the attenuation function allowing better noise reduction due to the better-ear effect.

The present article presents acoustic analysis and perceptual results as evidence that the binaural Fennec algorithm can improve speech reception in noise while preserving binaural cues. The perceptual results include measures of speech reception in noisy and reverberant conditions, as well as measures of sound source lateralization of target speech in the

presence of a masker. For the lateralization task, subjects were asked to detect if the target speech was coming from the left or the right side in the presence of background noise. All subjects could readily perform this lateralization task and clearly perceived the target speech as coming from either the left or the right; importantly, performance on this task was improved by the binaural Fennec algorithm when tested in an anechoic condition and was not significantly altered in the reverberant conditions. Consequently, the results indicate that the algorithm can improve speech reception in noise while preserving binaural cues necessary for lateralization.

II. METHODS

A. Subjects

Sixteen normal-hearing subjects participated in this study. The University of Southern California's Institute Review Board approved the study protocol. All subjects provided informed consent and were paid for their participation. All subjects were native English speakers who had pure tone audiometric thresholds of 20 dB hearing level or better at octave frequencies between 125 and 8000 Hz.

B. Materials

The coordinate response measure (CRM) sentence database (Bolia *et al.*, 2000) was used to measure speech reception and lateralization thresholds. The CRM materials consist of sentences of the form "Ready *callsign* go to *color number* now," with all 256 combinations of 8 call signs ("Arrow," "Baron," "Charlie," "Eagle," "Hopper," "Laker," "Ringo," "Tiger"), 4 colors ("blue," "green," "red," "white"), and 8 numbers (1–8). These sentence materials were recorded using four female and four male talkers, with an average sentence length of 3 s. For speech reception thresholds measured in the present study, only one of the talkers (a male) was used for both the target and the masker speech. The rationale for using the same talker on each trial is that in typical conversations one is aware of whom one is speaking to, while there are circumstances (e.g., answering the phone) that this assumption does not hold, it is typically true; therefore, we chose not to include talker variability as a perceptual dimension. In addition, for speech reception and lateralization testing, the competing talker masker was also selected as the same male speaker, but time-reversed. The rationale for using the same male talker for the competing masker was that we are primarily interested in energetic masking and wanted to minimize talker-specific cues that affect masking release such as vocal tract length and voicing cues.

Head-related transfer functions (HRTFs) to simulate spatial configurations were generated using a room simulation method known as the image method (Peterson, 1986; Shinn-Cunningham *et al.*, 2001). This simulation method was used since it provides a precise method for introducing certain aspects of interaural cues, such as interaural timing and level differences associated with head shadow, while precisely controlling reverberation levels. This simulation method does not simulate the effect of pinnae so does not capture spectral cues associated with elevation or front versus back asymmetries. This method does, however, provide precise control over inter-microphone placement and reverberation, which is useful for studies of spatial filtering since it allows primary factors of inter-microphone differences and reverberation levels to be examined while controlling for other factors such as measurement noise often encountered with measured HRTFs. The simulated room measured $4 \times 4 \times 2.6$ m with a 17-cm diameter reflective sphere located in the center of the room serving as a head model. Four microphone positions were rendered with two microphones on either side of the reflective sphere. The two microphones on either side of the head were separated by 1 cm in an endfire configuration (i.e., microphone array collinear with a target that is straight ahead of the listener). HRTFs were generated for sound source to microphone position for sound sources located 1 m away from the center of the sphere in the azimuthal plane for every angle from 0° to 360°, and for four different reverberation times (T60) including 0, 400, 800, 1200 ms. These HRTFs were used to spatialize target speech and masker speech in the various acoustic conditions described below for acoustic analysis and human subject testing.

C. Binaural Fennec algorithm

Goldsworthy *et al.* (2014) introduced a spatial filtering algorithm referred to as "Fennec" that uses two microphones situated 1 cm apart in an endfire configuration. The present article evaluates the performance of a binaural version of the Fennec algorithm, referred to as binaural Fennec. The binaural Fennec algorithm uses four microphone signals with two microphones over each ear in an endfire configuration. In this manner, there are four microphones total with two over each ear. To prevent the binaural version of the algorithm from producing any inherent interaural distortions, the left and right ear algorithms are combined to determine a joint spectrotemporal attenuation that is identically applied to the left and right ear signals. By applying the same spectrotemporal attenuation to both sides, the noise reduction processing emphasizes or suppresses specific spectrotemporal components without modifying interaural cues.

The Fennec algorithm compares the phase information of microphones that are 1 cm apart and situated over the ear. The first stage of Fennec processing is to compute a short-time Fourier transform of the front and back microphone signals. The implementation described in this article used a Fourier transform with 46.4-ms (1024-point) Hann windows with half-window overlap. Front and back microphone short-time Fourier transforms, $F(t,f)$ and $B(t,f)$, were used to calculate a spectrotemporal attenuation function based on inter-microphone phase differences.

Phase-based attenuation was calculated by estimating the angle of incidence (AOI) for each spectrotemporal component,

$$AOI(t,f) = \cos^{-1}\left(\frac{c}{d\pi f}\angle\frac{F(t,f)}{B(t,f)}\right), \quad (1)$$

where c is the speed of sound, d is the inter-microphone spacing, and $F(t,f)$ and $B(t,f)$ are the front and back microphone short-time Fourier transforms, respectively. The

estimated $AOI(t,f)$ is then transformed to an attenuation function,

$$A_{phase}(t,f) = \frac{N\left(AOI(t,f)|0,\beta\right)}{N(0|0,\beta)}, \quad (2)$$

where $N(x|\mu,\sigma)$ is the probability density function of a normal distribution with mean μ and standard deviation σ. In this equation, β is the beam width, which for the present evaluation was set to 30°. The resulting attenuation function has the shape of a normal probability density function but with a maximum value of 1 at 0°, and gradually approaching a value of 0 as the angle of incidence exceeds the beam width. Figure 1 illustrates the resulting attenuation as a function of the estimated angle of incidence.

The Fennec algorithm as described above was implemented on the left and right endfire pairs of microphones. Specifically, two different spectrotemporal weighting functions were calculated based on acoustic analysis of the left and right ear endfire pairs. At this point, one potential noise reduction solution would be to implement the Fennec algorithm independently on the left and right ear signals, but that solution would produce changes in the interaural characteristics. To avoid introducing such interaural distortions, it is necessary to apply the same spectrotemporal attenuation to the left and right ears. To accomplish this, the information from the left and right algorithms need to be combined to form a joint spectrotemporal attenuation.

There are multiple methods that might be used to combine information across ears, such as better ear analysis, but for this first examination of the binaural Fennec algorithm, a straightforward averaging of spectrotemporal attenuation is used. Specifically, spectrotemporal attenuation was calculated for the left and right ears using Eq. (2), then those two left and right attenuation functions were averaged to determine a joint spectrotemporal attenuation function [i.e., $A(t,f) = (A_L + A_R)/2$]. This attenuation function was then multiplied by the short-time Fourier transforms of the left and right back microphone signals. Since the same attenuation function is multiplied with the left and right microphone signals, any inherent inter-microphone timing and level differences will be preserved for each spectrotemporal component.

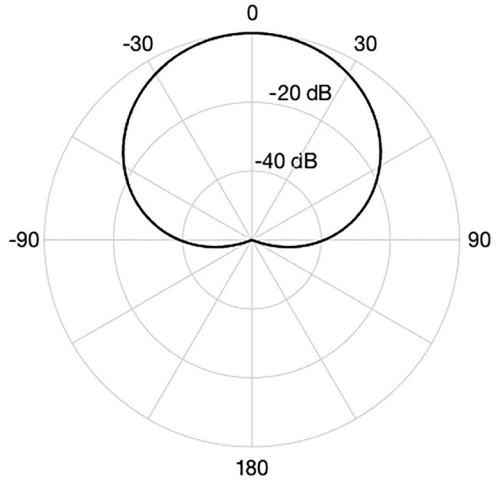

FIG. 1. Polar plot indicating attenuation as a function of angle of incidence for the Fennec algorithm as expressed in Eq. (2). For each spectrotemporal component (i.e., each time-frequency cell of the short-time Fourier transform), inter-microphone phase differences are used to estimate angle of incidence and then converted to an attenuation term per Eqs. (1) and (2).

The effect of this processing is to suppress spectrotemporal regions with poor target to masker ratios while preserving spectrotemporal regions with higher target to masker ratios. For the evaluations conducted here, the Binaural Fennec algorithm is implemented by jointly applying $A(t,f)$ to the left and right front microphone signals. This processed condition is always compared to an unprocessed condition which was simply routing the left and right front microphone signals to the left and right channels of the headphones.

D. Speech reception thresholds

Speech reception thresholds were measured for 8 conditions consisting of each combination of the four reverberation levels and with binaural Fennec processing compared to unprocessed conditions. The primary comparison in this measure is between binaural Fennec processed and unprocessed combined speech and noise. The speech reception procedure is always administered with the presence of the masker noise. Feedback was provided during the test with response buttons flashing green or red for correct and incorrect answers, respectively. A target sentence was randomly selected from the CRM database always using the same male talker. The target sentence was then filtered through the corresponding HRTF for 0°. The masker consisted of three randomly concatenated time-reverse sentences always using the same male talker as the target. The masker was then filtered through the corresponding HRTF for the 90° angle of incidence.

Subjects were instructed that the masker would be coming from 90° to the right but that they were listening for the sentence coming from straight ahead. The subjects were tested at four reverberation levels ($T_{60} = 0$, 400, 800, and 1200 ms). The target speech and masking time-reversed speech were processed using the same level of reverberation. The masking speech was time-reversed to reduce semantic cues as several studies have found substantial decrease in informational masking using time-reversed speech (Freyman *et al.*, 2001; Marrone *et al.*, 2008a; Iyer *et al.*, 2010; Best *et al.*, 2012; Gallun *et al.*, 2013; Swaminathan *et al.*, 2015; Kidd *et al.*, 2016).

The speech reception threshold procedure was implemented in MATLAB, the combined speech and noise was transmitted through an ESI U24XL external sound card, and presented to listeners at 65 dB sound pressure level through Sennheiser HD 280 pro headphones in a sound attenuating booth. All materials were down-sampled and processed to 16 000 Hz, but resampled to 44 100 Hz to avoid distortion effects associated with the digital-to-analog conversion. Subjects were unaware of any details associated with the acoustic and signal processing aspects of the conditions on which they were tested.

Sentences were scored correct when the subject identified both the color and number of the sentence. The initial target to masker ratio of the procedure was set to 0 dB, which was decreased/increased adaptively based on correct/incorrect answers. The step size of this decrease/increase started at 4 dB, multiplying by $2^{-1/4}$ for each reversal, until it reached a value of 2 dB on the fourth reversal and continued at that step size until the end of that run. The 1-up, 1-down procedure continued for eight reversals and the average target to masker ratio from the last four reversals was taken as the SRT for the run. The eight study conditions were tested in random order with three repetitions of each condition.

E. Lateralization thresholds

Lateralization thresholds were measured for the same eight conditions as the speech reception thresholds, but adjusting the target to masker ratio based on a subject's ability to correctly identify if a sentence was incident from the left or from the right of the straight-ahead direction. Like the speech reception procedure, the primary comparison in this measure is between binaural Fennec processed and unprocessed combined speech and noise. The lateralization is always performed with the presence of the masker. In this manner, the measured lateralization thresholds are an indicator of the noise tolerance for lateralization; specifically, measuring the minimum target to masker ratio at which the subjects could perform the lateralization task.

To measure lateralization thresholds, target speech was generated either 30° to the left or to the right of the straight-ahead direction and then combined with a time-reversed competing talker at 90° to the right as a masker. Based on the 8° minimum audible angle difference between a sound coming from 0° straight ahead, we proposed 30°, an angle greater than the minimum audible angle, to be sure that we would be measuring the noise tolerance of lateralization with and without the binaural Fennec algorithm, as opposed to tapping into their auditory perception limits (Carlile *et al.*, 2016).

Subjects were instructed to listen for the target speech, which was always the same sentence ("Ready Charlie, go to blue one now"), and determine if the target speech was coming from the left or from the right. All subjects reported that the target speech clearly appeared from a distinct spatial location from the time-reversed masker, and that they could easily perform the task until the target to masker ratio was substantially lowered. This task is relevant to listening situations such as when a person is attending to someone talking approximately straight ahead of them, but not precisely straight ahead. Being able to hear the lateralized position of the target speech may facilitate the listener's ability to attend to that speech; just as important, the ability to lateralize the target speech should contribute to the listener's sense of auditory space.

Lateralization thresholds were measured for the same eight conditions used for the speech reception thresholds, specifically the combinations of four reverberation levels ($T_{60} = 0$, 400, 800, and 1200 ms) with binaural Fennec compared to unprocessed conditions. The target sentence was filtered with an HRTF corresponding to either 30° to the left or to the right of the straight-ahead direction. Masking time-reversed speech was generated from a random selection of CRM sentences and filtered with an HRTF corresponding to 90° to the right of the straight-ahead direction. The average target to masker ratio across microphones was controlled by the adaptive procedure. The initial value of the target to masker ratio was 12 dB. The subject was asked the question "Is the sound coming from the left or the right?" in a two-alternative forced choice task.

The adaptive procedure was identical to the procedure used for speech reception thresholds with the exception that a 2-up, 1-down procedure was used, thus converging to 70.7% detection accuracy. This modification of was made since chance performance for the lateralization procedure was 50%, so a higher convergence point for detection accuracy was needed. In contrast, for the speech reception procedure chance performance for the speech reception procedure was only 3%, so the 1-up, 1-down procedure was sufficient. The lateralization procedure continued for eight reversals and the average target to masker ratio from the last four reversals was taken as the threshold for the run.

III. ACOUSTIC ANALYSES

Before examining the results of the speech reception and lateralization measures, acoustic analyses are presented to provide insight into algorithm performance. These acoustic analyses were implemented to quantify how well the binaural Fennec algorithm improves the target to masker ratio while preserving binaural cues. For these acoustic analyses, 16 different sentences were drawn from the CRM database for both the target and masker signals. The same talker was used for both the target and masker signal and the masker was time-reversed.

For the first acoustic analysis, the combined target plus masker signal was the anechoic condition with target at 0° and masker at 90° with 0 dB target to masker ratio. To analyze the effect of the binaural Fennec algorithm on interaural differences, the front left and front right microphone signals were compared in terms of inter-microphone timing and level differences. Figure 2 plots the inter-microphone timing and level differences between the left and right front microphones before and after binaural Fennec processing. The points represented in Fig. 2 are the individual spectrotemporal components (e.g., the individual time-frequency cells of the short-time Fourier transforms) of the signals, which are plotted by their individual timing and level differences. The solid line indicates the average energy associated with each inter-microphone timing or level difference. This average was calculated by summing component energy into bins using a histogram method and weighting the terms by the component energy. The upper left panel shows the distribution of signal power associated with time differences between microphones.

Two distribution clusters occur for interaural timing differences of 0 and 750 μs, corresponding to the target and masker locations of 0° and 90°, respectively. The left lower panel shows the corresponding distribution after binaural

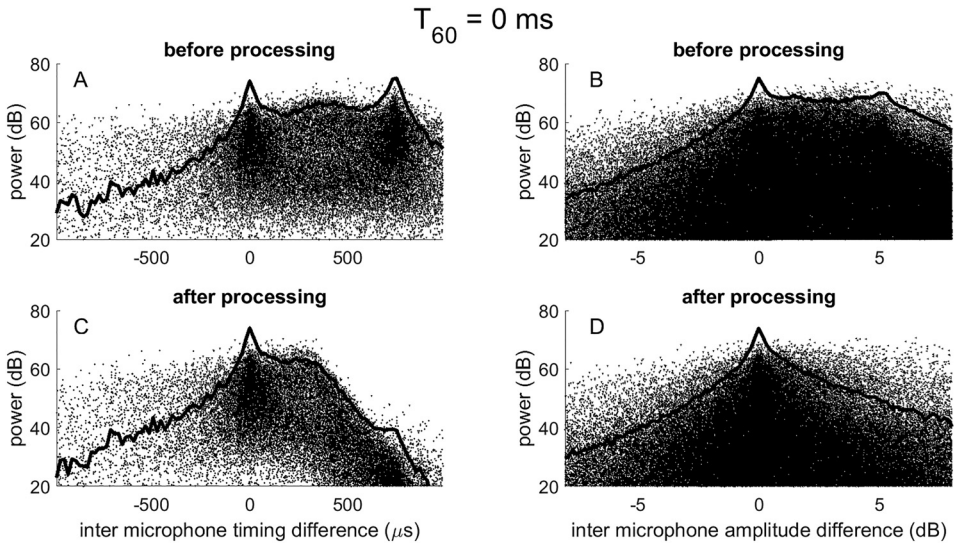

FIG. 2. Binaural Fennec preservation of binaural cues for the anechoic condition. Target speech was generated at 0° with a competing talker masker at 90°. The average target to masker ratio across microphones was 0 dB. Individual markers represent individual spectrotemporal components of the short-time Fourier transform as a power versus binaural cue pair. Solid line indicates a weighted histogram of power for these components. (A) The target at 0 us ITD and the masker at around 700 us ITD. After processing in (C), the masker power at 700 us is greatly attenuated, while the target power is generally preserved. (B) The target at 0 dB ILD and the masker at around 5 dB ILD. After processing in (C), the masker power at 5 dB is greatly attenuated, while the target power is generally preserved.

Fennec processing. The target distribution was relatively unchanged indicating that the time differences between microphones for the target speech were preserved while the energy associated with the masker speech was reduced. The timing differences associated with the masker speech were not changed, but the power associated with those components was reduced.

The inter-microphone level differences shown in the right panels of Fig. 2 have a distribution cluster at 0 dB corresponding to the target at 0° and with a wide spread distribution between 0 and 8 dB corresponding to the masker at 90°. The effect of binaural Fennec algorithm was again to emphasize the spectrotemporal components having inter-microphone time differences indicating the target signal. The processed inter-microphone level differences for the target and masker were not changed, but the power associated with those components was reduced.

The purpose of the preceding analysis was to substantiate the claim that the binaural Fennec algorithm preserves interaural timing and level cues. This claim is a straightforward consequence since the algorithm applies identical spectrotemporal attenuation to the left and right short-time Fourier transforms; consequently, the processing necessarily will preserve interaural differences. However, a relevant issue is the extent that such interaural cues are initially present in the left and right microphone signals prior to processing when the listener is in a reverberant environment. To investigate that issue, a second acoustic analysis was implemented using the same target and masker signals at 0 dB target to masker ratio, but considering a range of reverberation levels ($T_{60} = 0$, 400, 800, and 1200 ms).

For this second acoustic analysis, the spectrotemporal components of the short-time Fourier transform were divided into components that were either target or masker dominated based on whether the target to masker ratio for each spectrotemporal component was either greater or less than 0 dB. With this division of spectrotemporal components, a histogram analysis of interaural timing and level differences was calculated. Figure 3 illustrates equal-contour lines for the distributions of interaural timing and level differences associated with target and masker dominated components. The upper-left subplot indicates that the interaural distributions associated with the anechoic condition have distinct distributions for the target and masker dominated components.

This result provides acoustic differences for the Binaural Fennec algorithm to separate the target and masker components based on phase and/or level differences. The distinction between target and masker dominated components, however, becomes blurred with increasing reverberation. The upper-right subplot of Fig. 3 illustrates the distribution of interaural timing and level cues for the 400 ms reverberation time with much greater overlap in the target and masker dominated distributions. This blurring of interaural distributions worsens with increasing reverberation until there is complete overlap for the 1200 ms reverberation level.

A third acoustic analysis was completed to quantify the net improvement in the target to masker ratio produced by the algorithm for different reverberation levels ($T_{60} = 0$, 400, 800, and 1200 ms) when varying masker location in the azimuthal plane. The target speech was generated at a location of 0°, while the masker was generated at angles ranging from −180° to 180° in 30° increments. For each combination, the target and masker signals were filtered through corresponding HRTFs and added together such that the average target to masker ratio was 0 dB.

For these conditions, the combined target plus masker was filtered through the binaural Fennec algorithm, which produced a spectrotemporal attenuation function. In typical use, this spectrotemporal attenuation would be applied to the

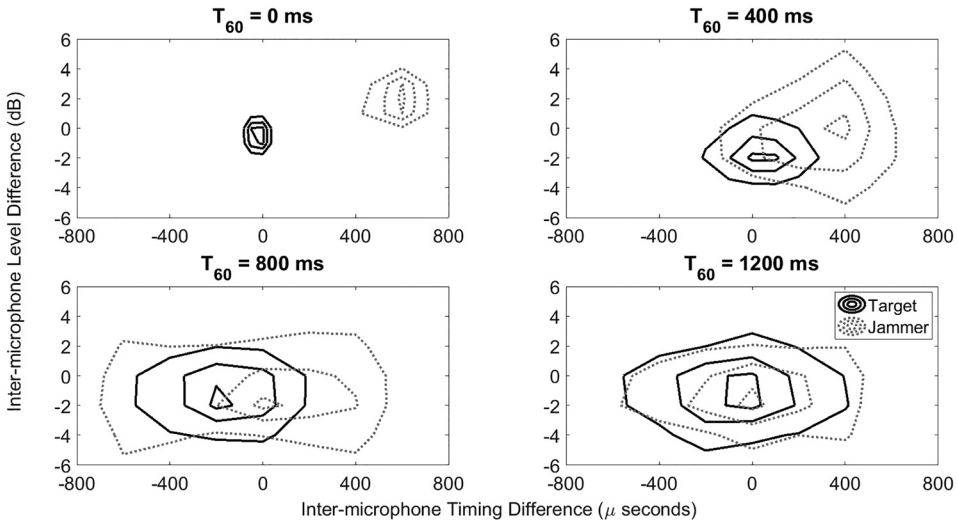

FIG. 3. Acoustic analysis of inter-microphone timing and level differences for simulated rooms having reverberation times (T_{60}) of 0, 400, 800, and 1200 ms.

short-time Fourier transform components of the left and right microphone signals before inverse transforming and presenting to the listener. For analysis purposes, this spectrotemporal attenuation was applied separately to the target and masker signals to quantify the net effects on the overall target to masker ratio. This type of algorithm analysis has been referred to as "yoked" processing (e.g., Greenberg and Zurek, 1992). Figure 4 illustrates the overall average attenuation of the left and right front microphone signals that results from the spectrotemporal attenuation applied separately to the target and masker.

Considering the anechoic condition, the algorithm did not suppress the masker when the masker was within the processing beam width (±30°). Once the masker was outside of the beam width, the masker was progressively attenuated reaching maximum attenuation near ±120°. Attenuation of the masker for a masker angle of 120° was approximately 17 dB, while attenuation of the target for that condition was approximately 1 dB, for a target to masker ratio improvement of approximately 16 dB. Performance was substantially degraded for the 400 ms condition with the primary effect being that the target was progressively attenuated. This effect was more pronounced for the 800 ms condition with more than 8 dB of target attenuation for all conditions. Consequently, for that condition, the overall improvement in the target to masker ratio was reduced to approximately 4 dB for conditions where the masker was outside of the defined beam width. While the observed target to masker benefit was

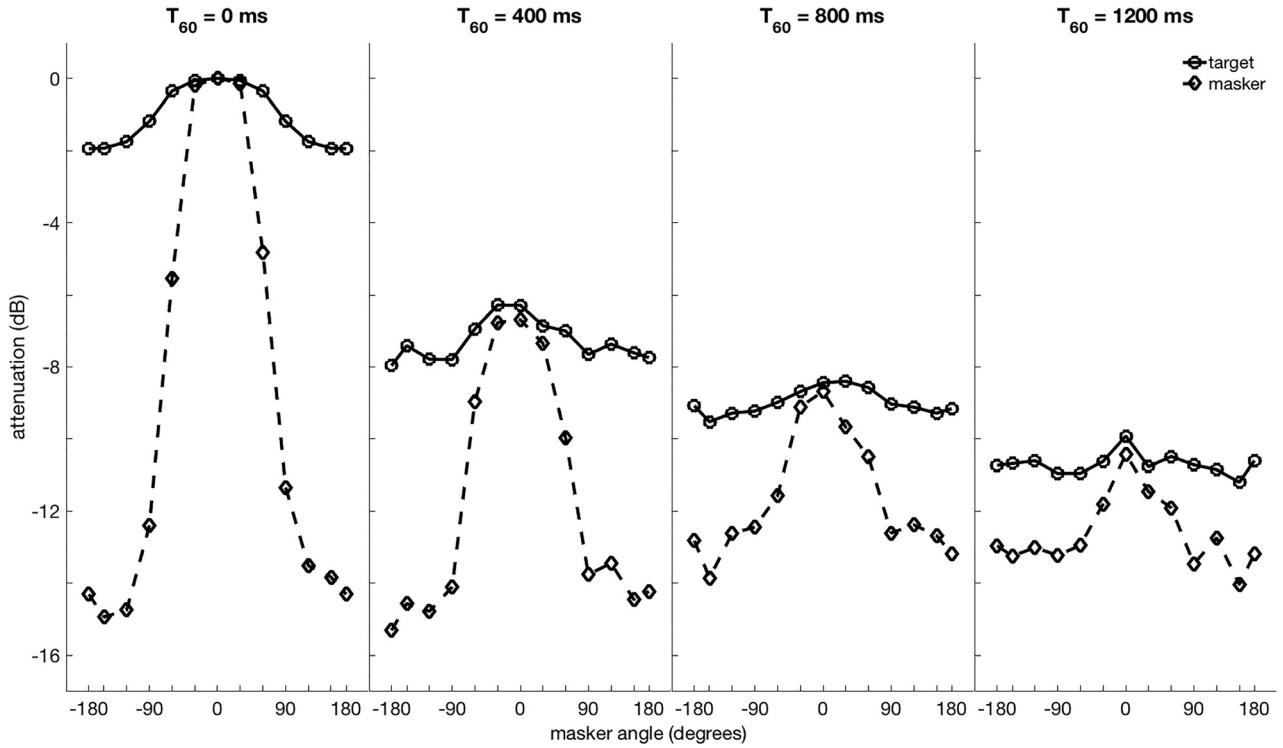

FIG. 4. Binaural Fennec attenuation of target and masker signals. Each panel shows results from the four reverberant conditions (T_{60} = 0, 400, 800, 1200 ms). The angle of incidence for the target speech was always 0°, results are plotted for masker angles ranging from −180° to 180° in 30° increments.

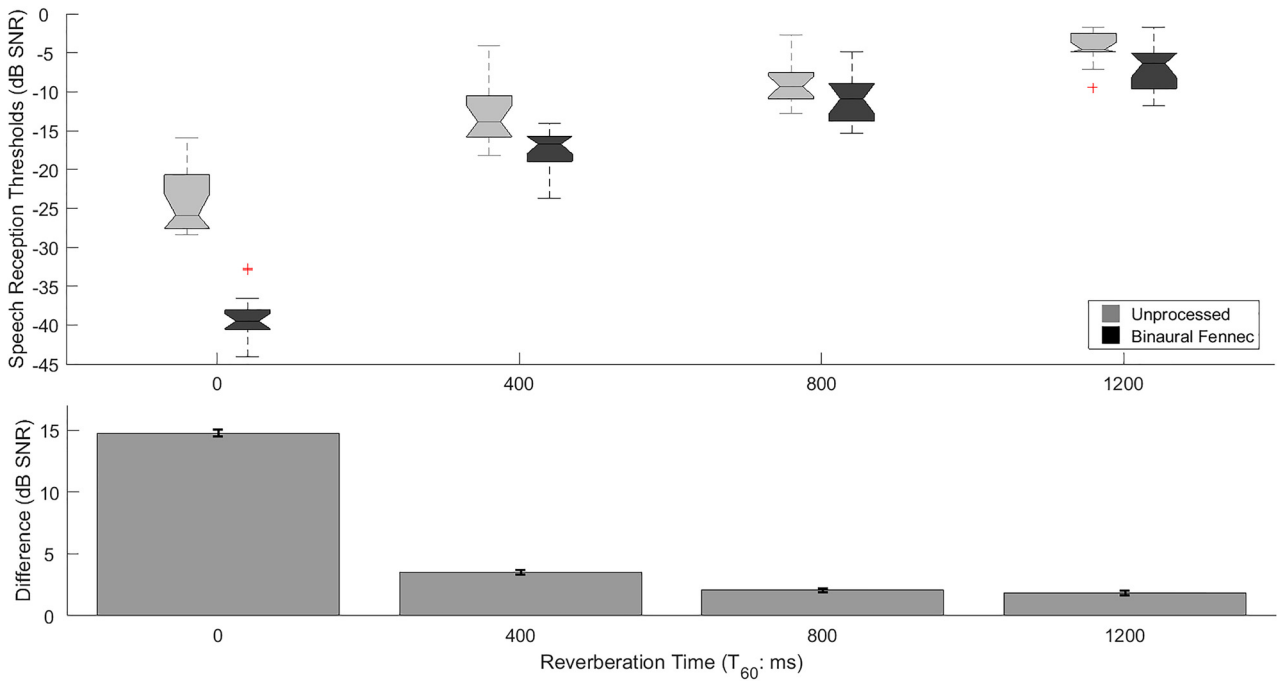

FIG. 5. (Color online) Speech reception thresholds (upper panel) and changes in speech reception thresholds derived from the Binaural Fennec algorithm (lower panel) for normal-hearing listeners with target speech at 0° and masker at 90° for each of the four reverberant conditions ($T_{60} = 0$, 400, 800, 1200 ms). Plotted speech reception thresholds are averages across repetitions and subjects, error bars indicate the standard error of the mean across subjects.

reduced compared to the less reverberant conditions, it was still a substantial benefit. For the 1200 ms condition, the effect was most pronounced causing more than 10 dB attenuation of the target, while the masker was attenuated by 13 dB.

The reason that algorithm performance degrades under increasing reverberation is that the inter-microphone timing differences used to estimate angle of incidence becomes increasingly diffuse. Target reflections combine with the direct component to indicate an angle of incidence that is outside of the processing beam, hence components are attenuated. While reverberation degrades the benefit derived from the algorithm, a consistent 3 to 4 dB improvement is still maintained for reverberant conditions.

IV. RESULTS

A. Speech reception thresholds

Speech reception thresholds were measured for eight conditions consisting of the combinations of four reverberation levels with the binaural Fennec algorithm compared to the unprocessed condition. The upper panel of Fig. 5 plots speech reception thresholds averaged across repetitions and subjects for each condition. The lower panel of Fig. 5 plots the average difference between speech reception thresholds between binaural Fennec and the unprocessed condition for each reverberation level. The general trend observed in these results is that speech reception thresholds increase with increasing levels of reverberation, but noting that binaural Fennec consistently improves thresholds compared to the unprocessed condition.

Analysis of variance (ANOVA) was implemented on the measured speech reception thresholds with reverberation level and algorithm (i.e., binaural Fennec or unprocessed) as factors and treating subject as a random blocking factor. Both reverberation level [$F(3,256) = 1948.8$, $p < 0.001$] and algorithm [$F(1,256) = 528.42$, $p < 0.001$] were significant effects confirming the clear trends observed in Fig. 5 that speech reception thresholds were degraded by reverberation and improved by the Binaural Fennec algorithm. The interaction between reverberation level and algorithm was significant [$F(3,256) = 149$, $p < 0.001$], reflecting the fact that the binaural Fennec algorithm provided more benefit in the anechoic condition, thus warranting *post hoc* analysis to determine if the observed speech reception benefits were significant at the other reverberation levels.

A *post hoc* multiple comparisons procedure was implemented on the statistics of the ANOVA described in the previous paragraph using Fisher's least significant difference method with a significance criterion of 0.05 to compare average speech reception thresholds with the binaural Fennec algorithm and the unprocessed conditions for each reverberation level. The comparisons at each reverberation level were significant. So while the speech reception benefits derived from the binaural Fennec algorithm decreased with increasing reverberation, the algorithm continued to provide significant benefits even for the highest level of reverberation tested.

B. Lateralization thresholds

Lateralization thresholds were measured for the same eight conditions as measured for speech reception, consisting of the combinations of four reverberation levels with the binaural Fennec algorithm compared to an unprocessed condition. The upper panel of Fig. 6 plots lateralization thresholds averaged across repetitions and subjects for each condition.

The lower panel of Fig. 6 plots the average difference between lateralization thresholds measured for the unprocessed conditions compared to binaural Fennec for each level of reverberation. The general trend is that lateralization thresholds increased with increasing levels of reverberation, noting that the binaural Fennec algorithm improved thresholds for the anechoic condition, but not for the other reverberant conditions.

An ANOVA was implemented for lateralization thresholds with reverberation level and algorithm (i.e., binaural Fennec or unprocessed) as factors and treating subject as a random blocking factor. Both reverberation [$F(3,256) = 523.4$, $p < 0.001$] and algorithm [$F(1,256) = 6.67$, $p = 0.0208$] were significant confirming the trend observed in Fig. 4 that lateralization thresholds were degraded by reverberation; however, the significance of algorithm is likely weighted by the exceptional performance of the binaural Fennec algorithm in the anechoic condition. This suspicion was substantiated in that the interaction between reverberation and algorithm was significant [$F(3,256) = 9.44$, $p < 0.001$], thus warranting *post hoc* analysis of algorithm comparisons at each reverberation level.

A *post hoc* multiple comparisons procedure was implemented on the statistics of the ANOVA described in the previous paragraph using Fisher's least significant difference method with a significance criterion of 0.05 to compare average lateralization thresholds with the binaural Fennec algorithm relative to the unprocessed conditions for each level of reverberation. With this criterion, the binaural Fennec algorithm only provided significant improvements in lateralization thresholds in the anechoic condition. The associated p values for comparing thresholds with the binaural Fennec algorithm and the unprocessed conditions at each reverberation level were $T_{60} = 0$ ms ($p < 0.001$), $T_{60} = 400$ ms ($p = 0.68$), $T_{60} = 800$ ms ($p = 0.99$), $T_{60} = 1200$ ms ($p = 0.55$). Stating this result in a positive manner, while the binaural Fennec algorithm did not improve lateralization thresholds in the reverberant conditions, it did not degrade them either. Thus, the results indicate that the algorithm improved speech reception thresholds while at least maintaining lateralization thresholds for the evaluated conditions.

C. Psychometric curve fitting

Additional analysis was conducted on the speech reception and lateralization thresholds to estimate psychometric curves for the observed data. For each acoustic condition, the detection accuracy (e.g., correct/incorrect response) was analyzed as a function of the target to masker ratio for every trial across subjects. A cumulative distribution function for a normal distribution was fit to this detection accuracy versus target to masker ratio comparison for each condition. Figure 7 illustrates the derived psychometric curves for each condition. The large benefits of the binaural Fennec algorithm for both speech reception and lateralization in the anechoic can be clearly seen in the upper subplots. For the reverberant conditions, the speech reception psychometric curves illustrate some improvement derived from the algorithm even in the most reverberant condition; however, no significant improvement was derived from the algorithm on the lateralization task. Consequently, for the present measures of speech reception and lateralization, the evidence indicates that the algorithm can improve speech reception while at least preserving lateralization of sound.

V. DISCUSSION

The binaural Fennec algorithm improved speech reception while preserving lateralization for normal-hearing

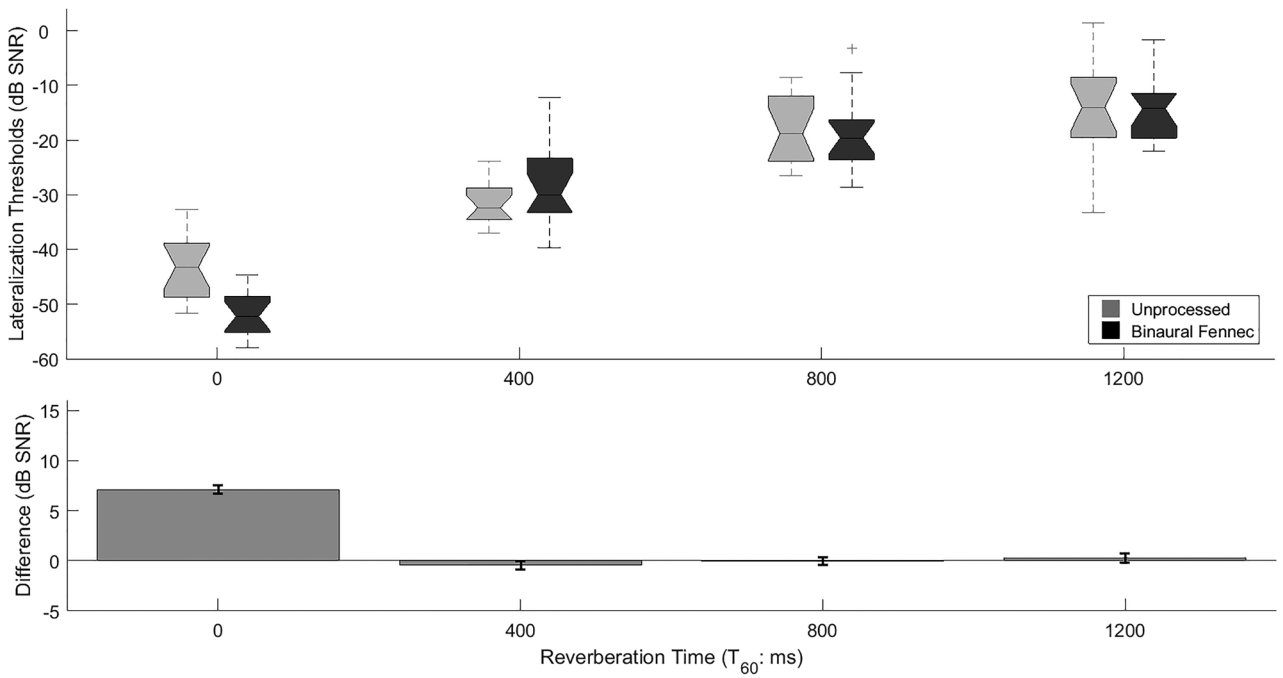

FIG. 6. Lateralization thresholds (upper panel) and changes in lateralization thresholds derived from the Binaural Fennec algorithm (lower panel) for normal-hearing listeners with target speech either 30° to the left or to the right and masker at 90° for each of the four reverberant conditions ($T_{60} = 0$, 400, 800, 1200 ms). Plotted lateralization thresholds are averages across repetitions and subjects, error bars indicate the standard error of the mean across subjects.

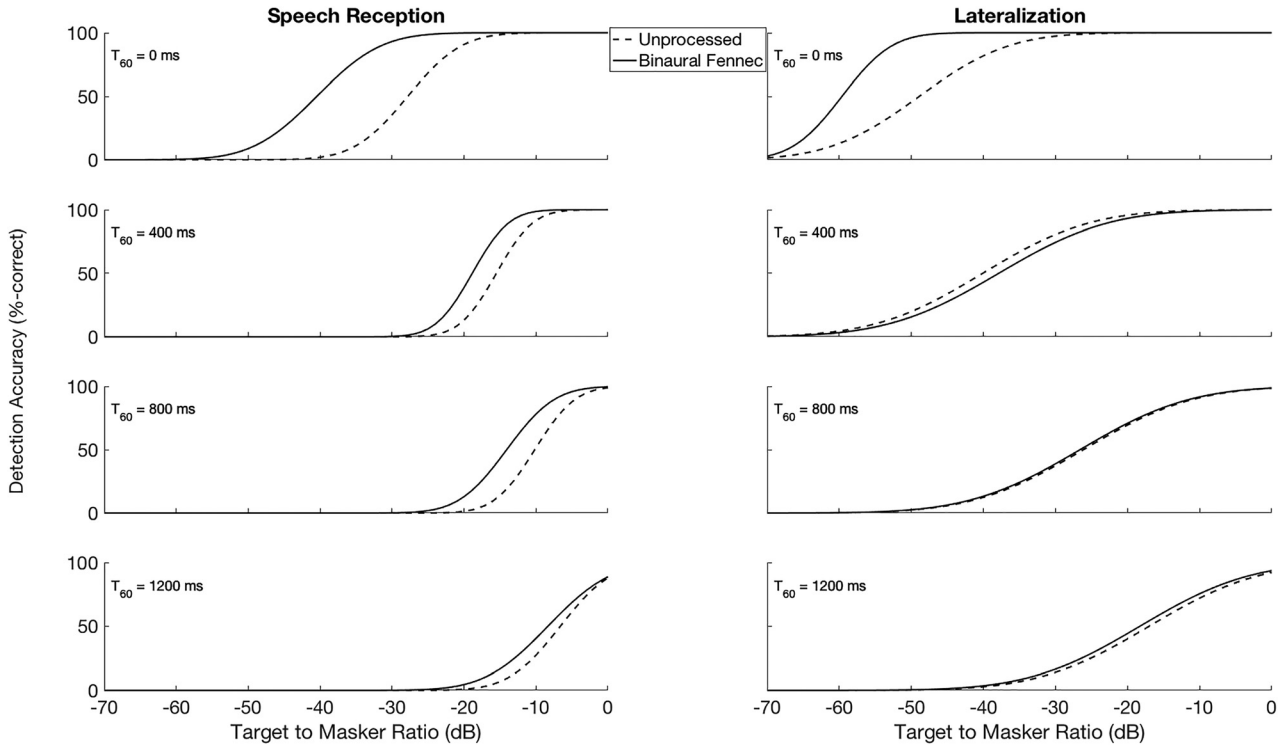

FIG. 7. Psychometric functions estimated from fitting cumulative distribution functions of Gaussian distributions to the measured detection accuracy for across all trials and subjects. Specifically, for every trial, detection accuracy was analyzed as a function of the target to masker ratio for that trial and the psychometric function was fitted by minimizing the mean-squared-error across every trial. The results indicate the Binaural Fennec algorithm improves speech reception thresholds, while preserving lateralization thresholds.

listeners for the conditions tested. This is an important result since adaptive spatial filtering algorithms have had mixed results for improving speech reception while preserving the interaural cues needed for sound localization. Certain modifications of beamforming algorithms avoided this issue by spectrally parsing the input signals and performing noise reduction or binaural cue preservation in different frequency regions, but not both (Welker *et al.,* 1997; Kidd *et al.,* 2015). Other approaches have used a trade-off function that allows a degree of binaural cue preservation to be maintained, but at the cost of noise suppression performance (Klasen *et al.,* 2006; Klasen *et al.,* 2007; Bogaert *et al.,* 2007; Bogaert *et al.,* 2008; Szurley *et al.,* 2016). The study described in the present article demonstrates that both objectives can be achieved, even in reverberant conditions.

There are several advantages for configuring spatial filtering for preserving binaural cues. Perhaps the most obvious is that a listener would be able to perceive the location of environmental sounds. This perceptual goal would be beneficial with respect to both the desired sound being attended to as well as any unwanted environmental sounds. It is important to note that while spatial filtering suppresses unwanted sounds, it generally does not completely remove those suppressed sounds. Consequently, the listener would still be able to hear sounds that are coming from locations removed from the target direction, but those sounds would simply be conveyed with less intensity. The procedures put forth by others to only preserve the binaural cues associated with target speech were not allowing the listener to associate a location, and consequently to orient, to other sounds in the environment.

Another advantage of preserving interaural timing and level cues is that it presumably would facilitate the listener's ability to segregate multiple sound sources into independent streams. A primary rationale for spatial filtering is the recognition that with hearing loss an individual's ability to spatially segregate sounds is diminished; consequently, spatial filtering is used to enhance sounds from a desired direction and, in a sense, perform the auditory stream segregation that occurs naturally in healthy auditory physiology. However, since hearing loss can have different degrees of severity, it would be better to provide spatial filtering preprocessing that retains binaural cues, which would allow the listener to make use of any residual stream segregation abilities that they retain. Individuals with mild hearing loss may have different attenuation versus interaural cue preservation trade-offs than individuals with severe to profound hearing loss.

This initial study of the binaural Fennec algorithm examined speech reception and lateralization thresholds in background noise for normal-hearing listeners. The motivation for using normal-hearing listeners was that normal-hearing listeners have relatively homogenous sound lateralization abilities. Specifically, the normal-hearing subjects could all perform the lateralization task and were impacted to a similar extent by reverberation and by additive background noise. Consequently, a clear baseline lateralization performance was determined and it was demonstrated that the binaural Fennec algorithm could preserve lateralization thresholds, and even improve lateralization thresholds for the anechoic condition. In this manner, using normal-hearing listeners for initial evaluations established an important

baseline for confirming that the binaural Fennec algorithm does not degrade lateralization thresholds.

This initial study of the binaural Fennec algorithm examined performance differences for the algorithm compared to an unprocessed condition, which was a simple routing of the left and right front microphone signals. This comparison provides a straightforward examination of algorithm performance that can be extrapolated to compare performance with other systems. For example, Chung *et al.* (2006) reported a 3.5 dB benefit for a hypercardioid directional over an omni-directional microphone. More generally, the overall improvement provided by directional relative to omni-directional microphones is roughly 3–5 dB in real-world environments with low reverberation compared to omni-directional microphones for listeners with acoustic hearing (Wouters *et al.,* 1999; Valente *et al.,* 2000; Amlani, 2001; Ricketts, 2001; Chung, 2004; Bentler, 2005). Comparing to this history of fixed directional microphones, the binaural Fennec algorithm outperforms fixed processing for anechoic and low levels of reverberation, but that the approaches yield similar results for more reverberant conditions.

As the field of adaptive spatial filtering evolves, it is important to clarify differences between emerging algorithms. The term "null-steering" has been traditionally used to describe the class of adaptive multiple microphone algorithms that are based on relatively slow (i.e., ∼500 ms) adaptations of a linear weighting of microphone signals to minimize the output noise power. This class of spatial filtering is appropriately referred to as null steering since an instantaneous view of the directional response of the system would indicate a spatial null that is steered towards the angle of incidence for the unwanted sound. So, if a masker existed at 90°, the null-steering beamformer would ideally place a null at 90° to cancel that noise. However, it would not simultaneously suppress other potential maskers at other angles. For a similar reason, we refer to the class of algorithms developed by Kollmeier and colleagues (Kollmeier *et al.,* 1993; Kollmeier and Koch, 1994) and expanded by others (Lockwood *et al.,* 2004; Goldsworthy *et al.,* 2014; Goldsworthy, 2014) as target-isolating beamformers since, rather than steer a spatial null, these algorithms make a spatial beam oriented at the target sound and suppress all other sounds.

This distinction between null-steering and target-isolating beamformers is relevant to their performance in reverberant environments. Although both classes of beamformers are affected by reverberation, there are differences regarding the mechanisms. Reverberation has two mechanisms of action on null-steering beamforming. First, reverberation causes the target speech to leak into the background noise estimator. This leakage reduces the performance of the calculations associated with minimizing the output noise power. Second, reverberation causes the background noise to become statistically spatially diffuse. Since null-steering beamformers are designed to orient a spatial null to cancel a single background noise source, they are not effective at cancelling diffuse noise. For target-isolating beamformers, the mechanism that reverberation affects performance is like the first mechanism described above. Specifically, with reverberation the target speech leaks outside of the algorithmic beam width and the target speech is attenuated. However, the second mechanism does not affect target-steering beamformers since this class of beamformer is designed to suppress sounds from all angles outside of the beam width. Consequently, the target-isolating beamformers are likely to be relatively robust to the fact that background noise becomes statistically diffuse with reverberation.

VI. CONCLUSION

Binaural Fennec is an adaptive multiple microphone spatial filtering algorithm designed to improve output target to masker ratio while preserving interaural timing and level differences. Results indicated that the binaural Fennec algorithm significantly improved speech reception thresholds in the presence of a competing talker at 90° even in reverberant conditions. Results also indicated that the binaural Fennec algorithm improved lateralization thresholds for an anechoic condition, while not significantly affecting threshold performance for reverberant conditions. This indicates that although binaural Fennec did not improve lateralization thresholds in reverberant conditions, it did not degrade the necessary spatial cues for lateralization. The conclusion drawn from these results is that the binaural Fennec algorithm can improve speech reception in noise, while preserving lateralization performance.

Amlani, A. M. (**2001**). "Efficacy of directional microphone hearing aids: A meta-analytic perspective," J. Am. Acad. Audiol. **12**, 202–214.

Arbogast, T. L., Mason, C. R., and Kidd, G., Jr. (**2005**). "The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **117**, 2169–2180.

Bentler, R. A. (**2005**). "Effectiveness of directional microphones and noise reduction schemes in hearing aids: A systematic review of the evidence," J. Am. Acad. Audiol. **16**, 473–484.

Berghe, J. V., and Wouters, J. (**1998**). "An adaptive noise canceller for hearing aids using two nearby microphones," J. Acoust. Soc. Am. **103**, 3621–3626.

Best, V., Marrone, N., Mason, C. R., and Kidd, G., Jr. (**2012**). "The influence of non-spatial factors on measures of spatial release from masking," J. Acoust. Soc. Am. **131**, 3103–3110.

Bogaert, T. V., Wouters, J., Doclo, S., and Moonen, M. (**2007**). "Binaural cue preservation for hearing aids using an interaural transfer function multichannel Wiener filter," IEEE Int. Conf. Acoust. Speech **4**, 565–568.

Bogaert, T. V. d., Doclo, S., Wouters, J., and Moonen, M. (**2008**). "The effect of multimicrophone noise reduction systems on sound source localization by users of binaural hearing aids," J. Acoust. Soc. Am. **124**(1), 484–497.

Bolia, R. S., Nelson, W. T., Ericson, M. A., and Simpson, B. D. (**2000**). "A speech corpus for multitalker communications research," J. Acoust. Soc. Am. **107**(2), 1065–1066.

Brandstein, M., and Ward, D. (**2001**). *Microphone Arrays: Signal Processing Techniques and Applications* (Springer, New York), pp. 88–105.

Brungart, D. S. (**2001**). "Informational and energetic masking effects in the perception of two simultaneous talkers," J. Acoust. Soc. Am. **109**(3), 1101–1109.

Carlile, S., Fox, A., Orchard-Mills, E., Leung, J., and Alais, D. (**2016**). "Six degrees of auditory spatial separation," J. Assoc. Res. Otolaryngol. **17**(3), 209–221.

Chung, K. (**2004**). "Challenges and recent developments in hearing aids. Part I: Speech understanding in noise, microphone technologies and noise reduction algorithms," Trends Amplif. **8**, 83–124.

Chung, K., and Zeng, F. G. (**2009**). "Using hearing aid adaptive directional microphones to enhance cochlear implant performance," Hear. Res. **250**, 27–37.
Chung, K., Zeng, F., and Acker, K. N. (**2006**). "Effects of directional microphone and adaptive multichannel noise reduction algorithm on cochlear implant performance," J. Acoust. Soc. Am. **120**(4), 2216–2227.
Chung, K., Zeng, F. G., and Waltzman, S. (**2004**). "Using hearing aid directional microphones and noise reduction algorithms to enhance cochlear implant performance," Acoust. Res. Lett. **5**, 56–61.
Desloge, J. G., Rabinowitz, W. M., and Zurek, P. M. (**1997**). "Microphone-array hearing aids with binaural output—Part I: Fixed-processing systems," IEEE Trans. Speech Audio. Process. **5**(6), 529–542.
Desmond, J. M., Collins, L. M., and Throckmorton, C. S. (**2014**). "The effects of reverberant self- and overlap-masking on speech recognition in cochlear implant listeners," J. Acoust. Soc. Am. **135**(6), EL304–EL310.
Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (**2001**). "Spatial release from informational masking in speech recognition," J. Acoust. Soc. Am. **109**, 2112–2122.
Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). "The role of perceived spatial separation in the unmasking of speech," J. Acoust. Soc. Am. **106**, 3578–3588.
Frost, O. L. (**1972**). "An algorithm for linearly constrained adaptive array processing," Proc. IEEE **60**(8), 926–935.
Gallun, F. J., Kampel, S. D., Diedesch, A. C., and Jakien, K. M. (**2013**). "Independent impacts of age and hearing loss on spatial release in a complex auditory environment," Front. Neurosci. **23**, 1–11.
Goldsworthy, R. L. (**2014**). "Two-microphone spatial filtering improves speech reception for cochlear-implant users in reverberant conditions with multiple noise sources," Trends Hear. **18**, 1–13.
Goldsworthy, R. L., Delhorne, L. A., Desloge, J. G., and Braida, L. D. (**2014**). "Two-microphone spatial filtering provides speech reception benefits for cochlear implant users in difficult acoustic environments," J. Acoust. Soc. Am. **136**(2), 867–876.
Greenberg, J. E., and Zurek, P. M. (**1992**). "Evaluation of an adaptive beamforming method for hearing aids," J. Acoust. Soc. Am. **91**, 1662–1676.
Griffiths, L. J., and Jim, C. W. (**1982**). "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. Antennas Propag. **30**, 27–34.
Hamacher, V., Doering, W. H., Mauer, G., Fleishmann, H., and Hennecke, J. (**1997**). "Evaluation of noise reduction systems for cochlear implant users in different acoustic environments," Am. J. Otol. **18**(6), S46–S49.
Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (**2004**). "The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer," J. Acoust. Soc. Am. **115**(2), 833–843.
Hazrati, O., and Loizou, P. C. (**2012**). "The combined effects of reverberation and noise on speech intelligibility by cochlear implant listeners," Int. J. Audiol. **51**(6), 437–443.
Hersbach, A. A., Arora, K., Mauger, S. J., and Dawson, P. W. (**2012**). "Combining directional microphone and single-channel noise reduction algorithms: A clinical evaluation in difficult listening conditions with cochlear implant users," Ear Hear. **33**(4), e13–e23.
Hirsh, I. J. (**1950**). "The masking of clicks by pure tones and bands of noise," J. Acoust. Soc. Am. **22**(5), 631–637.
Iyer, N., Brungart, D. S., and Simpson, B. D. (**2010**). "Effects of target-masker contextual similarity on the multimasker penalty in a three-talker diotic listening task," J. Acoust. Soc. Am. **128**, 2998–3010.
Jeffress, L. (**1948**). "A place theory of sound lateralization," J. Comput. Physiol. Psych. **41**, 35–39.
Kates, J. M. (**1993**). "Superdirective arrays for hearing aids," J. Acoust. Soc. Am. **94**, 1930–1933.
Kates, J. M., and Weiss, M. R. (**1996**). "A comparison of hearing-aid array processing techniques," J. Acoust. Soc. Am. **99**, 3138–3148.
Kidd, G., Jr., Mason, C. R., Best, V., and Marrone, N. (**2010**). "Stimulus factors influencing spatial release from speech-on-speech masking," J. Acoust. Soc. Am. **128**(4), 1965–1978.
Kidd, G., Jr., Mason, C. R., Best, V., and Swaminathan, J. (**2015**). "Benefits of acoustic beamforming for solving the cocktail party problem," Trends Hear. **19**, 1–15.
Kidd, G., Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., and Best, V. (**2016**). "Determining the energetic and informational components of speech-on-speech masking," J. Acoust. Soc. Am. **140**(1), 132–144.
Klasen, T., Doclo, S., Bogaert, T. V., Moonen, M., and Wouters, J. (**2006**). "Binaural multi-channel wiener filtering for hearing aids: Preserving interaural time and level differences," IEEE Int. Conf. Acoust. Speech **5**, 145–148.
Klasen, T. J., Van den Bogaert, T., Moonen, M., and Wouters, J. (**2007**). "Binaural noise reduction algorithms for hearing aids that preserve interaural time delay cues," IEEE Trans. Sign. Process. **55**(4), 1579–1585.
Kokkinakis, K., Azimi, B., Hu, Y., and Friedland, D. R. (**2012**). "Single and multiple microphone noise reduction strategies in cochlear implants," Trends Amplif. **16**, 102–116.
Kokkinakis, K., and Loizou, P. C. (**2010**). "Multi-microphone adaptive noise reduction strategies for coordinated stimulation in bilateral cochlear implant devices," J. Acoust. Soc. Am. **127**(5), 3136–3144.
Kollmeier, B., and Koch, R. (**1994**). "Speech enhancement based on physiological and psychoacoustical models of modulation perception and binaural interaction," J. Acoust. Soc. Am. **95**, 1593–1602.
Kollmeier, B., Peissig, J., and Hohmann, V. (**1993**). "Real-time multiband dynamic compression and noise reduction for binaural hearing aids," J. Rehabil. Res. Dev. **10**(1), 82–94.
Kompis, M., and Dillier, N. (**1994**). "Noise reduction for hearing aids: Combining directional microphones with an adaptive beamformer," J. Acoust. Soc. Am. **96**, 1910–1913.
Lockwood, M. E., Jones, D. L., Bilger, R. C., Lansing, C. R., O'Brien, W. D., Wheeler, B. C., and Feng, A. S. (**2004**). "Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms," J. Acoust. Soc. Am. **115**(1), 379–391.
Loizou, P., Hu, Y., Litovsky, R. Y., Yu, G., Peters, R., Lake, J., and Roland, P. (**2009**). "Speech recognition by bilateral cochlear implant users in a cocktail party setting," J. Acoust. Soc. Am. **125**, 372–383.
Marrone, N. L., Mason, C. R., and Kidd, G., Jr. (**2008a**). "Tuning in the spatial dimension: Evidence from a masked speech identification task," J. Acoust. Soc. Am. **124**, 1146–1158.
Marrone, N. L., Mason, C. R., and Kidd, G., Jr. (**2008b**). "Effect of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," J. Acoust. Soc. Am. **124**, 3064–3075.
Marrone, N. L., Mason, C. R., and Kidd, G., Jr. (**2008c**). "Evaluating the benefit of hearing aids in solving the cocktail party problem," Trends Amplif. **12**, 300–315.
Neher, T., Laugesen, S., Jensen, N. S., and Kragelund, L. (**2011**). "Can basic auditory and cognitive measures predict hearing impaired listeners' localization and spatial speech recognition abilities?," J. Acoust. Soc. Am. **130**, 1542–1558.
Peterson, P. M. (**1986**). "Simulating the response of multiple microphones to a single acoustic source in a reverberant room," J. Acoust. Soc. Am. **80**, 1527–1529.
Ricketts, T. A. (**2001**). "Directional hearing aids," Trends Amplif. **5**, 139–176.
Rothpletz, A. M., Wightman, F. L., and Kistler, D. J. (**2012**). "Informational masking and spatial hearing in listeners with and without unilateral hearing loss," J. Speech Lang. Hear. Res. **55**(2), 511–531.
Shinn-Cunningham, B. G., Desloge, J. G., and Kopco, N. (**2001**). "Empirical and modeled acoustic transfer functions in a simple room: Effects of distance and direction," in *Proceedings of the 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 183–186.
Soede, W., Berkhout, A. J., and Bilsen, F. A. (**1993a**). "Development of a directional hearing instrument based on array technology," J. Acoust. Soc. Am. **94**, 785–798.
Soede, W., Bilsen, F. A., and Berkhout, A. J. (**1993b**). "Assessment of a directional microphone array for hearing-impaired listeners," J. Acoust. Soc. Am. **94**, 799–808.
Spriet, A., Van Deun, L., Eftaxiadis, K., Laneau, J., Moonen, M., van Dijk, B., and Wouters, J. (**2007**). "Speech understanding in background noise with the two-microphone adaptive beamformer BEAM in the Nucleus Freedom Cochlear Implant System," Ear Hear. **28**(1), 62–72.
Stadler, R. W., and Rabinowitz, W. M. (**1993**). "On the potential of fixed arrays for hearing aids," J. Acoust. Soc. Am. **94**, 1332–1342.
Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V. A., Kidd, G., Jr., and Patel, A. D. (**2015**). "Musical training, individual differences and the cocktail party problem," Sci. Rep. **5**, 1–10.
Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Roverud, E., and Kidd, J. G. (**2016**). "Role of binaural temporal fine structure and envelope cues in cocktail-party listening," J. Neurosci. **36**(31), 8250–8257.
Szurley, J., Bertrand, A., Dijk, B. V., and Moonen, M. (**2016**). "Binaural noise cue preservation in a binaural noise reduction system with a remote microphone signal," IEEE Trans. Audio Speech **24**(5), 952–966.

Thiemann, J., Müller, M., Marquardt, D., Doclo, S., and Par, S. V. (**2016**). "Speech enhancement for multimicrophone binaural hearing aids aiming to preserve the spatial auditory scene," EURASIP J. Adv. Sign. Process. **2016**(1), 1–11.
Valente, M., Schuchman, G., Potts, L. G., and Beck, L. B. (**2000**). "Performance of dual-microphone in-the-ear hearing aids," J. Am. Acad. Audiol. **11**, 181–189.
Van Hoesel, R. J., and Clark, G. M. (**1995**). "Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients," J. Acoust. Soc. Am. **97**(4), 2498–2503.
Van Veen, B. D., and Buckley, K. M. (**1988**). "Beamforming: A versatile approach to spatial filtering," IEEE ASSP Mag. **5**(2), 4–24.
Welker, D. P., Greenberg, J. E., Desloge, J. G., and Zurek, P. M. (**1997**). "Microphone-array hearing aids with binaural output—Part II: A two-microphone adaptive system," IEEE Trans. Speech Audio Process. **5**, 543–551.
Woods, W. S., Kalluri, S., Pentony, S., and Nooraei, N. (**2013**). "Predicting the effect of hearing loss and audibility on amplified speech reception in a multi-talker listening scenario," J. Acoust. Soc. Am. **133**, 4268–4278.
Wouters, J., Litere, L., and VanWieringen, A. (**1999**). "Speech intelligibility in noisy environments with one and two microphone hearing aids," Audiology **38**, 91–98.
Wouters, J., and Vanden Berghe, J. (**2001**). "Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system," Ear Hear. **22**, 420–430.