Comparing the effects of reverberation and of noise on speech recognition in simulated electric-acoustic listening

Kate Helms Tillery, Christopher A. Brown,^{a)} and Sid P. Bacon

*Psychoacoustics Laboratory, Department of Speech and Hearing Science, Arizona State University,
P.O. Box 870102, Tempe, Arizona 85287-0102*

(Received 8 September 2010; revised 28 October 2011; accepted 7 November 2011)

Cochlear implant users report difficulty understanding speech in both noisy and reverberant environments. Electric-acoustic stimulation (EAS) is known to improve speech intelligibility in noise. However, little is known about the potential benefits of EAS in reverberation, or about how such benefits relate to those observed in noise. The present study used EAS simulations to examine these questions. Sentences were convolved with impulse responses from a model of a room whose estimated reverberation times were varied from 0 to 1 sec. These reverberated stimuli were then vocoded to simulate electric stimulation, or presented as a combination of vocoder plus low-pass filtered speech to simulate EAS. Monaural sentence recognition scores were measured in two conditions: reverberated speech and speech in a reverberated noise. The long-term spectrum and amplitude modulations of the noise were equated to the reverberant energy, allowing a comparison of the effects of the interferer (speech vs noise). Results indicate that, at least in simulation, (1) EAS provides significant benefit in reverberation; (2) the benefits of EAS in reverberation may be underestimated by those in a comparable noise; and (3) the EAS benefit in reverberation likely arises from partially preserved cues in this background accessible via the low-frequency acoustic component.

© 2012 Acoustical Society of America. [DOI: 10.1121/1.3664101]

PACS number(s): 43.66.Ts, 43.71.Ky [RYL]

Pages: 416–423

I. INTRODUCTION

Current cochlear implant (CI) technology allows many deafened individuals to recover a significant degree of auditory speech recognition. While outcomes vary, average sentence intelligibility scores can exceed 80% correct in quiet environments (Wilson and Dorman, 2008). Performance in competing backgrounds, however, is substantially poorer (e.g., Skinner *et al.*, 2002; Nelson *et al.*, 2003; Spahr *et al.*, 2007; Kokkinakis and Loizou, 2011). For example, fluctuating noise at a signal-to-noise ratio (S/N) of +8 dB can cause a 50 percentage-point drop in intelligibility for CI users,¹ compared to performance in quiet (Nelson *et al.*, 2003).

Another type of “noise” that poses a significant problem for CI users is reverberation. In self-report measures, reverberant environments are rated as being equal to noise in difficulty (Plyler *et al.*, 2008; Donaldson *et al.*, 2009). Reverberant energy smears the amplitude envelope, prolongs phoneme durations, obscures voicing onsets and offsets, and blurs formant transitions (see Brown and Palomaki, 2006, for a review). CI users find these temporal distortions particularly deleterious because it is chiefly temporal cues upon which they rely for speech reception (Poissant *et al.*, 2006; Kokkinakis and Loizou, 2011). The intelligibility of CI-processed speech can drop by more than 60 percentage points when reverberation time [RT₆₀; American National Standards Institute (ANSI), 2002] increases from 0 to 1000 ms (Poissant *et al.*, 2006; Kokkinakis and Loizou, 2011).

Evidence from studies of speech reception by normal-hearing listeners in acoustic spaces indicates that, while discrete early reflections assist intelligibility (Soulodre *et al.*, 1989), densely spaced later reflections effectively mask the direct signal (Lochner and Burger, 1964). These masking effects have been likened to those of noise (Bolt and MacDonald, 1949), which has led to the use of S/N-type metrics to predict speech intelligibility in reverberation (e.g., Lochner and Burger, 1961, 1964; Houtgast and Steeneken, 1973). It has also been suggested that performance in noise that has been equated to reverberant energy may predict performance in reverberation (Gelfand and Silman, 1979), a notion revisited in the present study.

For CI users, the difficulties imposed by complex backgrounds such as noise and reverberation may both arise from a limited access to fine-structure cues (Moore, 2003; Qin and Oxenham, 2003, 2005). Standard CI signal-processing strategies discard the fine structure of complex signals and convey only envelope modulations across a few spectral bands. While an interdependence of fine structure and envelope make it difficult to isolate one from the other (Ghitza, 2001; Heinz and Swaminathan, 2009; Apoux *et al.*, 2011), researchers generally agree that envelope cues are sufficient for speech reception in quiet, given adequate spectral detail (e.g., Lorenzi *et al.*, 2006; Hopkins *et al.*, 2008; Gilbert and Lorenzi, 2010). In noisy environments, however, fine-structure cues are necessary (Lorenzi *et al.*, 2006), perhaps because they convey information about pitch (Smith *et al.*, 2002; Brown and Bacon, 2010).

Voice pitch cues such as fundamental frequency (F0) variation and differences in mean F0 are important for good speech reception, both in noise (Brown and Bacon, 2010)

^{a)}Author to whom correspondence should be addressed. Electronic mail: c-b@asu.edu

and in reverberation (Darwin and Hukin, 2000). Fine-structure processing may be integral to the use of these cues, as it plays a role in such tasks as the perception of pitch for harmonic complexes (Moore, 2008b) and the detection of frequency modulation (Moore and Sek, 1995). Standard CI processors do convey weak pitch cues about voice gender (Fu *et al.*, 2005) and intonation patterns (Chatterjee and Peng, 2008). However, this degree of fine-structure sensitivity does not correlate with speech reception performance in complex backgrounds (Drennan *et al.*, 2008).

CI simulations using a vocoder have shown that increasing fine-structure detail by adding frequency modulation to the already amplitude-modulated carriers can improve speech reception in noise (Stickney *et al.*, 2005) and reverberation (Drgas and Blaszkak, 2010). While these improvements in simulation are intriguing, there remain considerable practical difficulties associated with increasing fine-structure detail for CI users. Ongoing efforts to increase fine structure via the creation of “virtual” electrodes (e.g., McDermott and McKay, 1994; Geurts and Wouters, 2004; Firszt *et al.*, 2007; Berenstein *et al.*, 2008; Landsberger and Srinivasan, 2009) or the provision of fine temporal detail (e.g., Rubinstein *et al.*, 1999; Arnoldner *et al.*, 2007; Riss *et al.*, 2009) have met with limited success. In general, results show some improvement in pitch detection or discrimination, but limited gains to speech reception in noise (Berenstein *et al.*, 2008; Firszt *et al.*, 2009; Han *et al.*, 2009; Riss *et al.*, 2009; Shannon *et al.*, 2011).

For CI users with residual acoustic hearing, another way to increase fine-structure cues involves the combined use of acoustic and electric stimulation. Referred to variably as electric-acoustic stimulation (EAS) or bimodal hearing, this configuration involves the use of residual low-frequency acoustic hearing in the ear either ipsilateral or contralateral to the CI (Gifford *et al.*, 2010). The improvements to intelligibility in noise with this strategy are well-documented (e.g., von Ilberg *et al.*, 1999; Gantz and Turner, 2004; Gantz *et al.*, 2005; Kong *et al.*, 2005; Gifford *et al.*, 2007; Dorman *et al.*, 2008; Brown and Bacon, 2009a; Zhang *et al.*, 2010). Dorman *et al.* (2008) reported that, in a babble background at a S/N of +5 dB, CI users’ sentence recognition improved by an average of 22 percentage points with the addition of contralateral acoustic stimulation. The majority of the EAS benefit is present when the acoustic region is limited to frequencies below 125 Hz (for a male target talker) (Zhang *et al.*, 2010), or when it is replaced by a pure tone, modulated to track changes in amplitude and in F0 of the target speech (Brown and Bacon, 2009a,b). This suggests that voice pitch cues are likely the source of EAS benefit in noise (Brown and Bacon, 2009a,b, 2010).

Many CI users with residual hearing report an EAS benefit in reverberation as well as noise, citing better speech reception and more natural sound (Fitzpatrick *et al.*, 2009). Aside from subjective reports such as these, little is known quantitatively about speech reception with EAS in reverberant environments. Given that the acoustic component of EAS improves access to F0 cues (Brown and Bacon, 2009a, 2009b), and that these cues benefit normal-hearing listeners in reverberation (Darwin and Hukin, 2000), CI users with residual hearing might also obtain a benefit. The first aim of

the present study was to explore the potential benefits of combining electric and low-frequency acoustic stimulation in reverberation using EAS simulations.

The degrees to which noise and reverberation affect speech reception in EAS is equally unclear. These two backgrounds might reduce intelligibility by similar amounts, as has been argued for listeners with normal hearing (Bolt and MacDonald, 1949; Lochner and Burger, 1961, 1964; Gelfand and Silman, 1979). On the other hand, reverberation might be less detrimental than noise for EAS users. While the temporal envelope of speech is smeared in reverberation, fine-structure cues such as F0 variation are preserved to some degree (Darwin and Hukin, 2000). Access to such cues via acoustic stimulation may mitigate the effects of reverberation. The acoustic component presumably would not be as helpful in a comparable noise, however, because F0 cues in the target speech would be masked energetically by the background noise.

The second aim of the present study was to examine the relationship between the effects of reverberation and those of noise on speech intelligibility in simulated EAS, by comparing performance in reverberation with that in a noise that had been equated to the reverberation along several acoustic dimensions. The rationale for this manipulation arose from a pilot study in which the levels of reverberation and speech-shaped noise had been adjusted to yield similar vocoder-only intelligibility scores. We found that adding a 500-Hz low-pass component (simulating EAS) resulted in larger benefits in reverberation than in noise. Our recent identification of F0 variation as an important cue in EAS (Brown and Bacon, 2009a,b) led us to hypothesize that the pilot results had arisen due to the energetic masking of F0 cues in noise, but the partial preservation of these cues in reverberation. However, because the two backgrounds had not been equated acoustically, it was difficult to draw such a conclusion from these data. To address this in the present study, we sought to generate a reverberated noise whose long-term spectrum and amplitude envelope were equated to the reverberation, but whose fine structure differed from that of the reverberant energy. This allowed a direct comparison of the effects of the fine structure in each background.

The present study aimed to examine the effects of reverberation on speech intelligibility in simulated electric and electric-acoustic stimulation, and compare them with those in an equated noise. We predicted that in vocoder-alone conditions, the effects of reverberated speech and reverberated noise would be similar, as vocoder processing would more or less discard fine-structure cues. We further predicted that in simulated EAS, greater benefit would be observed in reverberation than in reverberated noise, because F0 variation would be relatively preserved in reverberation, but would be masked in the equated noise.

II. METHODS

A. Subjects

Twenty students from Arizona State University participated. Hearing screenings confirmed pure-tone thresholds in both ears of 20 dB HL or better at octave and half-octave frequencies from 125 to 6000 Hz. Participants ranged in age

from 18 to 24 years (mean = 19.7 years) and were native speakers of American English. All participants provided informed consent and were paid \$10 per hour.

B. Stimuli

Target stimuli were 400 pre-recorded sentences, taken from the Institute of Electrical and Electronics Engineers (IEEE) corpus (IEEE, 1969). Each sentence contained five key words and was produced by a male speaker. The sentences were combined with three backgrounds: anechoic, simulated reverberation, and a noise with characteristics resembling those of the reverberant energy. The target-plus-background mixtures were then processed to simulate electric stimulation, low-frequency acoustic stimulation, and EAS.

C. Signal processing

1. Simulated reverberation

Reverberation was simulated using a MATLAB implementation of an image-based room acoustics model (McGovern, 2004). We modeled a rectangular space the size of a small classroom (4.59 m wide $\times$ 6.64 m long $\times$ 2.6 m high). The relative position of the sound source was $1.43 \times 6.25 \times 1.3$ m, and the position of the receiver in the room was $2.8 \times 2.5 \times 1.3$ m (see Fig. 1). The model assumes that the reflective characteristics of the room's six surfaces are uniform and invariant with respect to frequency. The model also assumes an empty room. That is, it does not account for volume or size of the source or receiver, only position. The absorption coefficient of the room's surfaces was varied to simulate estimated reverberation times (RT_{60}) of 0, 125, 250, 500, and 1000 ms. Impulse responses were derived from this model, and convolved with each speech token to simulate reverberant speech. In the case of an RT of 0 ms, the convolution step was omitted.

2. Reverberated noise

A noise with a temporal envelope and a long-term spectrum comparable to those of the reverberant energy was

FIG. 1. Schematic showing the relative positions of the sound source and the receiver in a reverberant room (4.59 m $\times$ 6.64 m $\times$ 2.6 m). Source: 1.43 m $\times$ 6.25 m $\times$ 1.3 m; 0° azimuth from the center of the receiver. Receiver: 2.8 m $\times$ 2.5 m $\times$ 1.3 m.

generated using the following procedure. The long-term average spectrum of broadband Gaussian white noise was shaped to match that of the IEEE sentences. The amplitude envelope of each IEEE sentence, extracted via half-wave rectification and low-pass filtering (16 Hz, second-order Butterworth filter), was used to modulate the speech-shaped noise. The rms level of the noise was equated with the rms level of the target sentence. Simulated reverberation was then applied to the amplitude modulated speech-shaped noise using the impulse responses described above. The direct, anechoic portion of the impulse response was removed from the reverberated impulse response by subtraction. Finally, the reverberated noise was combined with its corresponding IEEE sentence. As in the reverberated speech background, the convolution step was omitted when RT was 0 ms.

3. Cochlear implant simulation

Electric listening was simulated using a four-channel sinusoidal vocoder we have used in the past (Brown and Bacon, 2009b). For each sentence, the target-plus-background mixture was band-pass filtered into four adjacent frequency bands with logarithmically spaced cut-off frequencies of 750, 1234, 2031, 3342, and 5500 Hz. The amplitude envelope of each band was extracted via half-wave rectification and low-pass filtering (sixth-order Butterworth filters with a low-pass cut-off of either 400 Hz or half the bandwidth of the frequency band, whichever was lower). Each envelope was applied to a tone at the arithmetic center of each band. The frequencies of the tones were 992, 1633, 2687, and 4421 Hz. The resulting modulated tones were summed to represent simulated electric hearing.

4. EAS simulation

Low-frequency residual hearing was simulated by low-pass filtering the target-plus-background mixture at 500 Hz (tenth-order Butterworth). Electric-acoustic hearing was simulated by combining the vocoder with the low-pass filtered speech.

D. Procedures

All processing was accomplished digitally using MATLAB. The stimuli were output via an Echo Gina3G sound card through a Tucker-Davis PA5 programmable attenuator, and presented monaurally through a Sennheiser HD250 II headphone. On average, the playback level was 73 dB SPL for the broadband speech, and 70 dB SPL for the vocoder-plus-low-pass speech. On average, the level of the vocoder was within 1–2 dB of the level of the low-pass speech.

Intelligibility scores for unprocessed, low-pass filtered, vocoded, and vocoded-plus-low-pass-filtered speech were measured at each RT in reverberation and in reverberated noise. The five RTs (0, 125, 250, 500, and 1000 ms), four processing conditions (broadband speech, vocoder, 500-Hz low-pass speech, and vocoder plus 500-Hz low-pass speech), and two backgrounds (reverberated speech and reverberated noise) resulted in a total of 40 conditions. Half of the participants heard speech in reverberation first, and half heard

speech in reverberated noise first. Within each background, the order of conditions was randomized for each participant.

All testing was done in a double-walled sound-attenuated booth. Prior to testing, participants were familiarized with the speech material by listening to 20 IEEE sentences (10 unprocessed, 10 vocoded)² presented at an overall level of 70 dB SPL. Participants were instructed to repeat each sentence as completely as possible. No feedback was provided. The number of key words correct per sentence was recorded by the experimenter. Each condition consisted of 10 sentences, resulting in a total of 50 key words per condition. The raw score out of 50 key words was converted to percent correct.

III. RESULTS

Figure 2 shows mean percent correct speech intelligibility as a function of reverberation time, in a background of reverberation [Fig. 2(a)] and reverberated noise [Fig. 2(b)], with the different plots representing the four different processing conditions. Three of the processing conditions yielded

FIG. 2. Mean percent correct speech intelligibility as a function of reverberation time for a background of (a) reverberation and (b) reverberated noise. Diamonds represent broadband stimuli; filled triangles represent vocoder plus 500-Hz low-pass filtered speech; squares represent vocoder; unfilled triangles represent 500-Hz low-pass filtered speech. Error bars represent ± 1 standard error.

similar patterns of results across the two backgrounds. Intelligibility of broadband speech remained between 80% and 100% correct across all RTs in both backgrounds, although it was slightly lower in reverberated noise than in reverberated speech. Intelligibility of vocoded speech alone dropped by more than 30 percentage points with even a relatively short RT of 125 ms, and remained low across the longer RTs. Low-pass filtered speech alone yielded scores around 10% correct or lower in both backgrounds (i.e., $RT_{60} \geq 125$ ms), although they were slightly lower in reverberated noise.

On the other hand, the processing condition simulating EAS produced performance in reverberation that was different from that in reverberated noise. The combination of 500-Hz low-pass speech with vocoded speech (V/500; filled triangles) produced substantially higher scores (over vocoder alone) in reverberation than in reverberated noise. Across all RTs (0–1000 ms), the average improvement over vocoder alone provided by the addition of the 500-Hz low-pass component was 41 percentage points in reverberation, compared to nine percentage points in reverberated noise.

A three-factor between-subjects analysis of variance showed all interactions were significant ($p < 0.001$). Main effects of background ($p < 0.001$), processing condition ($p < 0.001$), and reverberation time ($p < 0.001$) were also significant. A *post hoc* Tukey analysis with a Bonferroni correction was performed, targeting six pairwise comparisons. Mean speech intelligibility scores were not significantly different across background for the vocoder condition ($p > 0.05$), although they were different for broadband speech ($p < 0.001$) and 500 Hz low-pass speech ($p = 0.01$). Mean speech intelligibility scores were significantly different across background for the V/500 processing condition ($p < 0.001$). Mean performance with V/500 was significantly higher than with vocoder alone in reverberated speech ($p < 0.001$), but not in reverberated noise ($p > 0.05$).

IV. DISCUSSION

The present study examined the benefits of combining electric and acoustic stimulation in reverberation and compared them to those in a corresponding noise. While there exist methods for equating reverberation to other types of noise (Lochner and Burger, 1961, 1964; Houtgast and Steeneken, 1973; Houtgast *et al.*, 1980), the motivation underlying the present study necessitated a different approach. In this case, equating the long-term spectrum and envelope of the noise to those of the reverberation allowed an examination of the effects of differences in fine structure across the two backgrounds.

It was hypothesized that the reverberated noise and the reverberated speech would degrade intelligibility to similar degrees in simulated electric stimulation, since fine structure is largely discarded by the vocoder, but would affect intelligibility differently in simulated EAS. It was further hypothesized that simulated EAS would ameliorate some of the deleterious effects of reverberation by providing access to fine-structure cues, but would not provide such a benefit in noise, because of energetic masking effects.

The pattern of results supports these hypotheses. In reverberation and in reverberated noise, broadband speech remains mostly intelligible, even at longer reverberation times. Listeners' recognition of vocoded speech, however, is substantially degraded in both backgrounds: scores drop considerably at even relatively short RTs. The addition of low-pass filtered speech to the vocoded signal restores a substantial degree of intelligibility in reverberation, even at longer RTs. Conversely, in reverberated noise, the vocoder/low-pass combination provides negligible benefit. The main difference between the two backgrounds was the presence of fine-structure cues. Fine structure from the target speech was present (albeit smeared) in the reverberated speech background, but absent from the reverberated noise background. These results suggest that fine-structure cues accessible via the low-pass component are likely a source of benefit for EAS users in reverberation.

These results also suggest that noise, when carefully equated to reverberation, may be a reasonable predictor of CI performance in reverberant environments, at least in simulation. This is because conventional CI processing essentially eliminates any fine-structure information preserved by reverberation, and yields performance levels that are similar to those achieved in noise. In simulated EAS, however, even carefully equated noise may not provide a good indication of performance in reverberation. This is likely because the low-frequency acoustic component of EAS provides access to the fine-structure cues that are preserved in reverberation, and thus yields performance levels that are higher than those achieved in noise that would mask these cues.

One candidate for the particular fine-structure cue responsible for the EAS benefit in reverberation is F0 variation. Partially preserved F0 cues in reverberated speech have been observed to influence intelligibility for normal-hearing listeners. Culling *et al.* (2003) found that reverberated sentences ($RT_{60} = 400$ ms) with inverted F0 contours had a deleterious effect on intelligibility when a background talker was present. Darwin and Hukin (2000) found that the natural prosody of speech (e.g., F0 contours and word stress) helped listeners selectively attend to sentences in reverberation ($RT_{60} = 400$ ms). Even when background talkers are absent, as in the present study, it may be the case that F0 cues help listeners track the target amid the "noise" of temporally dense later reflections.

Consideration of F0 as the primary source of EAS benefit in reverberation does not preclude a potential additional contribution from formant frequencies, particularly F1 cues. Given the filtering parameters used to simulate residual low-frequency hearing in the present study, it is possible that the first formants of some vowels were at times audible (Peterson and Barney, 1952; Hillenbrand *et al.*, 1995). Studies are underway to examine the individual contributions of F0 and formant frequencies to the EAS benefit in reverberation.

Consideration should also be given to a possible contribution of early reflections to the EAS benefit in reverberation. Whitmal and Poissant (2009) have noted that CI processing reduces the ability typically seen in normal-hearing listeners to fuse early reflections with the direct signal. In simulated electric-only listening, their subjects did

not benefit from fusion, and in some cases showed a decrement in performance. In the present study, it is possible that partial access to this benefit was restored when low-pass filtered speech was combined with the vocoder. Speech intelligibility in listeners with high-frequency hearing loss has been shown to withstand the addition of a single reflection at delay times up to 40 ms (Nábělek and Robinette, 1978), suggesting that a similar outcome might occur in simulated EAS, given access to fine-structure cues in the low-pass speech. The fine structure in discrete early reflections of the direct signal, accessible via the low-pass component, may provide some benefit in the reverberated speech, but would be absent in the reverberated noise. Studies are underway to examine the effect of early reflections in EAS, including the potential contribution of particular fine-structure cues such as F0.

While temporal fine-structure mechanisms are generally viewed as responsible for the coding of F0 (Moore, 2008a,b), spectral resolution also plays a role in pitch tasks (e.g., Moore, 2003). Thus, it is possible that spectral cues may have contributed to the present results. However, given the estimated widths of auditory filters at low frequencies (Moore and Glasberg, 1983) and the degree of F0 variation observed in typical speech (Traunmüller and Eriksson, 2000), it is unlikely that F0 ranged widely enough to result in a large contribution from spectral cues.

Two technical issues warrant mention. One concerns the nature of the reverberated noise. Because one aim of the present study was a direct comparison of the effects of noise and of reverberation in simulated EAS, the direct-to-reverberant ratio was kept similar across the two backgrounds. The level of the target speech compared to the level of the masker in reverberated noise was thus relatively low (e.g., -8.6 dB at an RT of 1000 ms). While the signal-to-noise or direct-to-reverberant ratio was not a variable of interest *per se*, it should be acknowledged that the level of the noise likely contributed to the low performance with simulated EAS in the reverberated noise condition (the average benefit above vocoder-only was about nine percentage points). These results should not be taken to imply, however, that EAS provides little benefit in noise. On the contrary, a very large EAS benefit is observed in more "typical" noise backgrounds, and at more typical signal-to-noise ratios. For example, in their study of EAS in noise, Zhang *et al.* (2010) found an average EAS benefit over electric-only stimulation of 47 percentage points in multi-talker babble at a signal-to-noise ratio of $+10$ dB.

In addition to the possible effects of level, the smearing of the envelope modulations in the reverberated noise may have provided fewer opportunities for listeners to glimpse the target in the dips of the masker, compared to the noise used in Zhang *et al.* (2010). These factors may also have contributed to the slightly lower performance with broadband and low-pass filtered speech in this background. An interesting follow-up experiment would involve examining speech intelligibility using the noise in its anechoic form (prior to reverberation). This essentially speech-modulated noise might facilitate the separation of energetic masking effects (which should be similar for both the reverberated

and non-reverberated noises) and temporal effects, such as smearing of the amplitude envelope.

The nature of the vocoder also merits comment. Two considerations motivated the use of sinusoidal carriers, rather than noise bands. First, the percept provided by sinusoidal vocoders is reportedly closer to that received by actual CI users, in terms of its tonal quality (Dorman *et al.*, 1997). Second, although fine-structure cues in vocoded speech are relatively poor regardless of the type of carrier used, sinusoidal vocoders have been shown to transmit more fine-structure information than those with noise band carriers (Whitmal *et al.*, 2007; Stone *et al.*, 2008). Had any useful fine-structure cues been present in the vocoded signal, our sinusoidal vocoder would have provided the better opportunity for listeners to access them. Vocoder-only performance was about 20 percentage points greater at 125 ms in reverberation than in reverberated noise, suggesting that some fine-structure was available. Nevertheless, our results in simulation support the notion that EAS can improve speech intelligibility in reverberation, and that fine-structure cues serve as the primary source of benefit.

While caution must be used in applying conclusions drawn from simulations to actual EAS users, our findings do suggest that CI users with residual hearing may experience an EAS benefit in reverberant environments: the addition of low-frequency acoustic hearing to their electric stimulation may provide a substantial boost to speech intelligibility. Further, this benefit may be greater than would be predicted from the EAS benefit in a comparable noise background.

The potential benefit from fine-structure cues in reverberation might arguably be limited by hearing loss in the low-frequency acoustic region and the resulting reduction in fine-structure processing (Lorenzi *et al.*, 2006). However, individuals with cochlear hearing loss are often reported to retain some fine-structure processing, depending on the task and the degree of hearing loss (e.g., Buss *et al.*, 2004; Hopkins and Moore, 2007; Santurette and Dau, 2007; Hopkins *et al.*, 2008; Ardoint *et al.*, 2010). For example, Buss *et al.* (2004) observed that, given a 500-Hz carrier, some listeners with moderate hearing loss demonstrated frequency modulation detection thresholds that were 1–2 Hz away from the normal range. Additionally, Hopkins and colleagues found that the ability to use fine-structure cues in a sentence recognition task was affected by hearing loss more at high than at low frequencies (Hopkins and Moore, 2007; Hopkins *et al.*, 2008). This suggests that, even when low-frequency hearing is impaired, there still exists some ability to benefit from fine-structure cues when listening to speech in noise. Further, fine-structure-based pitch perception, even using impaired low-frequency acoustic hearing, is arguably superior to that obtained from conventional electric stimulation. Evidence for this arises from the well-documented improvements to speech reception in noise when low-frequency acoustic stimulation is added to electric stimulation (von Ilberg *et al.*, 1999; Gantz and Turner, 2004; Gantz *et al.*, 2005; Kong *et al.*, 2005; Gifford *et al.*, 2007; Dorman *et al.*, 2008; Brown and Bacon, 2009a; Zhang *et al.*, 2010).

Evidence is emerging that EAS does indeed provide a benefit in reverberation. Data in our laboratory from six EAS

users show that, in a reverberation time of 500 ms, intelligibility scores with EAS were on average 32 percentage points higher than those obtained with electric-only stimulation. Research continues into the benefits seen by actual EAS users in reverberation, versus those seen in noise. In addition, efforts are ongoing to shed light on the specific acoustic cues and auditory processes underlying this benefit.

V. SUMMARY

The intelligibility of vocoded speech in reverberation increases significantly when low-pass filtered speech is added to the signal. The benefits of combining low-pass speech with the vocoder are greater in reverberation than in a noise whose long-term spectrum and amplitude envelope are comparable to those of the reverberant energy. F0 cues that are partially preserved in the reverberant speech and accessible via the low-pass component are likely responsible for the benefit. These results suggest that CI users who are able to make use of residual acoustic hearing will improve their speech intelligibility in reverberant environments, and may do so to a greater degree than would be predicted by their performance in a comparable noise.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute on Deafness and Other Communication Disorders (NIDCD Grant No. DC008329). The authors gratefully acknowledge William A. Yost for many helpful discussions about this work.

¹To maintain clarity, hearing-impaired individuals who do not use a cochlear implant will be referred to as “individuals with hearing loss,” and hearing-impaired individuals who use a cochlear implant will be referred to as “CI users.”

²As noted in previously published work from this laboratory (Brown and Bacon, 2009b) vocoder-only performance after practice with 400 sentences was no different from that observed after practice with 10 sentences. To limit the potential effects of fatigue, we opted to keep the training as brief as possible.

ANSI (2002). *Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools* (American National Standards Institute, New York), Pub. no. S12.60-2002.

Apoux, F., Millman, R. E., Viemeister, N. F., Brown, C. A., and Bacon, S. P. (2011). “On the mechanisms involved in the recovery of envelope information from temporal fine structure,” *J. Acoust. Soc. Am.* **130**, 273–282.

Ardoint, M., Sheft, S., Fleuriot, P., Garnier, S., and Lorenzi, C. (2010). “Perception of temporal fine-structure cues in speech with minimal envelope cues for listeners with mild-to-moderate hearing loss,” *Int. J. Audiol.* **49**, 823–831.

Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W.-D., and Hamzavi, J.-S. (2007). “Speech and music perception with the new fine structure speech coding strategy: Preliminary results,” *Acta Oto-Laryngol.* **127**, 1298–1303.

Berenstein, C. K., Mens, L. H. M., Mulder, J. J. S., and Vanpoucke, F. J. (2008). “Current steering and current focusing in cochlear implants: Comparison of monopolar, tripolar, and virtual channel electrode configurations,” *Ear Hear.* **29**, 250–260.

Bolt, R. H., and MacDonald, A. D. (1949). “Theory of speech masking by reverberation,” *J. Acoust. Soc. Am.* **21**, 577–580.

Brown, C. A., and Bacon, S. P. (2009a). “Achieving electric-acoustic benefit with a modulated tone,” *Ear Hear.* **30**, 489–493.

Brown, C. A., and Bacon, S. P. (2009b). “Low-frequency speech cues and simulated electric-acoustic hearing,” *J. Acoust. Soc. Am.* **125**, 1658–1665.

- Brown, C. A., and Bacon, S. P. (2010). "Fundamental frequency and speech intelligibility in background noise," *Hear. Res.* **266**, 52–59.
- Brown, G. J., and Palomaki, K. J. (2006). "Reverberation," in *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, edited by D. Wang and G. J. Brown (Wiley Interscience, Hoboken, NJ), Chap. 7, pp. 209–250.
- Buss, E., Hall, J. W., and Grose, J. H. (2004). "Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss," *Ear Hear.* **25**, 242–250.
- Chatterjee, M., and Peng, S.-C. (2008). "Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition," *Hear. Res.* **235**, 143–156.
- Culling, J. F., Hodder, K. I., and Toh, C. Y. (2003). "Effects of reverberation on perceptual segregation of competing voices," *J. Acoust. Soc. Am.* **114**, 2871–2876.
- Darwin, C. J., and Hukin, R. W. (2000). "Effects of reverberation on spatial, prosodic, and vocal-tract size cues to selective attention," *J. Acoust. Soc. Am.* **108**, 335–342.
- Donaldson, G. S., Chisolm, T. H., Blasco, G. P., Shinnick, L. J., Ketter, K. J., and Krause, J. C. (2009). "BKB-SIN and ANL predict perceived communication ability in cochlear implant users," *Ear Hear.* **30**, 401–410.
- Dorman, M. F., Gifford, R. H., Spahr, A. J., and McKarns, S. A. (2008). "The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies," *Audiol. Neurotol.* **13**, 105–112.
- Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band output," *J. Acoust. Soc. Am.* **102**, 2409–2411.
- Drennan, W. R., Longnion, J. K., Ruffin, C., and Rubinstein, J. T. (2008). "Discrimination of Schroeder-phase harmonic complexes by normal-hearing and cochlear-implant listeners," *JARO* **9**, 138–149.
- Drgas, S., and Blaszkak, M. A. (2010). "Perception of speech in reverberant conditions using AM- FM cochlear implant simulation," *Hear. Res.* **269**, 162–168.
- Firszt, J. B., Holden, L. K., Reeder, R. M., and Skinner, M. W. (2009). "Speech recognition in cochlear implant recipients: Comparison of standard HiRes and HiRes 120 sound processing," *Otol. Neurotol.* **30**, 146–152.
- Firszt, J. B., Koch, D. B., Downing, M., and Litvak, L. (2007). "Current steering creates additional pitch percepts in adult cochlear implant recipients," *Otol. Neurotol.* **28**, 629–636.
- Fitzpatrick, E. M., Séguin, C., Schramm, D., Chenier, J., and Armstrong, S. (2009). "Users' experience of a cochlear implant combined with a hearing aid," *Int. J. Audiol.* **48**, 172–182.
- Fu, Q.-J., Chinchilla, S., Nogaki, G., and Galvin, J. J. (2005). "Voice gender identification by cochlear implant users: The role of spectral and temporal resolution," *J. Acoust. Soc. Am.* **118**, 1711–1718.
- Gantz, B. J., and Turner, C. W. (2004). "Combining acoustic and electrical speech processing: Iowa/Nucleus hybrid implant," *Acta Oto-Laryngol.* **124**, 344–347.
- Gantz, B. J., Turner, C. W., Gfeller, K. E., and Lowder, M. W. (2005). "Preservation of hearing in cochlear implant surgery: Advantages of combined electrical and acoustical speech processing," *Laryngoscope* **115**, 796–802.
- Gelfand, S. A., and Silman, S. (1979). "Effects of small room reverberation upon the recognition of some consonant features," *J. Acoust. Soc. Am.* **66**, 22–29.
- Geurts, L., and Wouters, J. (2004). "Better place-coding of the fundamental frequency in cochlear implants," *J. Acoust. Soc. Am.* **115**, 844–852.
- Ghitza, O. (2001). "On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception," *J. Acoust. Soc. Am.* **110**, 1628–1640.
- Gifford, R. H., Dorman, M. F., McKarns, S. A., and Spahr, A. J. (2007). "Combined electric and contralateral acoustic hearing: Word and sentence recognition with bimodal hearing," *J. Speech Lang. Hear. Res.* **50**, 835–843.
- Gifford, R. H., Dorman, M. F., Shallop, J. K., and Sydlowski, S. A. (2010). "Evidence for the expansion of adult cochlear implant candidacy," *Ear Hear.* **31**, 186–194.
- Gilbert, G., and Lorenzi, C. (2010). "Role of spectral and temporal cues in restoring missing speech information," *J. Acoust. Soc. Am.* **128**, 294–299.
- Han, D., Liu, B., Zhou, N., Chen, X., Kong, Y., Liu, H., Zheng, Y., and Xu, L. (2009). "Lexical tone perception with HiResolution and HiResolution 120 sound-processing strategies in pediatric Mandarin-speaking cochlear implant users," *Ear Hear.* **30**, 169–177.
- Heinz, M. G., and Swaminathan, J. (2009). "Quantifying envelope and fine-structure coding in auditory nerve responses to chimaeric speech," *JARO* **10**, 407–423.
- Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099–3111.
- Hopkins, K., and Moore, B. C. J. (2007). "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information," *J. Acoust. Soc. Am.* **122**, 1055–1068.
- Hopkins, K., Moore, B. C. J., and Stone, M. A. (2008). "Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech," *J. Acoust. Soc. Am.* **123**, 1140–1153.
- Houtgast, T., and Steeneken, H. J. M. (1973). "The modulation transfer function in room acoustics as a predictor of speech intelligibility," *Acustica* **28**, 66–73.
- Houtgast, T., Steeneken, H. J. M., and Plomp, R. (1980). "Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acoustics," *Acustica* **46**, 59–72.
- Institute of Electrical and Electronics Engineers (IEEE) (1969). "IEEE recommended practice for speech quality measurements," *IEEE Trans. Audio. Electroacoust.* **17**, 225–246.
- Kokkinakis, K., and Loizou, P. C. (2011). "The impact of reverberant self-masking and overlap-masking effects on speech intelligibility by cochlear implant listeners," *J. Acoust. Soc. Am.* **130**, 1099–1102.
- Kong, Y.-Y., Stickney, G. S., and Zeng, F.-G. (2005). "Speech and melody recognition in binaurally combined acoustic and electric hearing," *J. Acoust. Soc. Am.* **117**, 1351–1361.
- Landsberger, D. M., and Srinivasan, A. G. (2009). "Virtual channel discrimination is improved by current focusing in cochlear implant recipients," *Hear. Res.* **254**, 34–41.
- Lochner, J. P. A., and Burger, J. F. (1961). "The intelligibility of speech under reverberant conditions," *Acustica* **11**, 195–200.
- Lochner, J. P. A., and Burger, J. F. (1964). "The influence of reflections on auditorium acoustics," *J. Sound Vib.* **1**, 426–454.
- Lorenzi, C., Gilbert, G., Cam, H., Garnier, S., and Moore, B. C. J. (2006). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *Proc. Natl. Acad. Sci. USA* **103**, 18866–18869.
- McDermott, H. J., and McKay, C. M. (1994). "Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea," *J. Acoust. Soc. Am.* **96**, 155–162.
- McGovern, S. G. (2004). "A model for room acoustics," <http://www.2pi.us/rir.html> (Last viewed October 27, 2011).
- Moore, B. C., and Sek, A. (1995). "Effects of carrier frequency, modulation rate, and modulation waveform on the detection of modulation and the discrimination of modulation type (amplitude modulation versus frequency modulation)," *J. Acoust. Soc. Am.* **97**, 2468–2478.
- Moore, B. C. J. (2003). "Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants," *Otol. Neurotol.* **24**, 243–254.
- Moore, B. C. J. (2008a). "The role of temporal fine structure in normal and impaired hearing," in *Auditory Signal Processing in Hearing-impaired Listeners, 1st International Symposium on Auditory and Audiological Research (ISAAR 2007)*, edited by T. Dau, J. M. Buchholz, J. M. Harte, and T. U. Christiansen (Centertryk A/S, Denmark).
- Moore, B. C. J. (2008b). "The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people," *JARO* **9**, 399–406.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Nábělek, A. K., and Robinette, L. (1978). "Influence of the precedence effect on word identification by normally hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.* **63**, 187–194.
- Nelson, P. B., Jin, S.-H., Carney, A. E., and Nelson, D. A. (2003). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," *J. Acoust. Soc. Am.* **113**, 961–968.
- Peterson, G. E., and Barney, H. L. (1952). "Control methods used in a study of the vowels," *J. Acoust. Soc. Am.* **24**, 175–184.
- Plyler, P. N., Bahng, J., and von Hapsburg, D. (2008). "The acceptance of background noise in adult cochlear implant users," *J. Speech Lang. Hear. Res.* **51**, 502–515.

- Poissant, S. F., Whitmal, N. A., and Freyman, R. L. (2006). "Effects of reverberation and masking on speech intelligibility in cochlear implant simulations," *J. Acoust. Soc. Am.* **119**, 1606–1615.
- Qin, M. K., and Oxenham, A. J. (2003). "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," *J. Acoust. Soc. Am.* **114**, 446–454.
- Qin, M. K., and Oxenham, A. J. (2005). "Effects of envelope-vocoder processing on F0 discrimination and concurrent-vowel identification," *Ear Hear.* **26**, 451–460.
- Riss, D., Arnoldner, C., Reiß, S., Baumgartner, W.-D., and Hamzavi, J.-S. (2009). "1-year results using the Opus speech processor with the fine structure speech coding strategy," *Acta Oto-Laryngol.* **129**, 988–991.
- Rubinstein, J. T., Wilson, B. S., Finley, C. C., and Abbas, P. J. (1999). "Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation," *Hear. Res.* **127**, 108–118.
- Santurette, S., and Dau, T. (2007). "Binaural pitch perception in normal-hearing and hearing-impaired listeners," *Hear. Res.* **223**, 29–47.
- Shannon, R. V., Cruz, R. J., and Galvin, J. J., 3rd (2011). "Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise," *Audiol. Neurotol.* **16**, 113–123.
- Skinner, M. W., Arndt, P. L., and Staller, S. J. (2002). "Nucleus 24 advanced encoder conversion study: Performance versus preference," *Ear Hear.* **23**, 2S–17S.
- Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). "Chimaeric sounds reveal dichotomies in auditory perception," *Nature* **416**, 87–90.
- Soulodre, G. A., Popplewell, N., and Bradley, J. S. (1989). "Combined effects of early reflections and background noise on speech intelligibility," *J. Sound Vib.* **135**, 123–133.
- Spahr, A. J., Dorman, M. F., and Loisel, L. H. (2007). "Performance of patients using different cochlear implant systems: Effects of input dynamic range," *Ear Hear.* **28**, 260–275.
- Stickney, G. S., Nie, K., and Zeng, F.-G. (2005). "Contribution of frequency modulation to speech recognition in noise," *J. Acoust. Soc. Am.* **118**, 2412–2420.
- Stone, M. A., Füllgrabe, C., and Moore, B. C. J. (2008). "Benefit of high-rate envelope cues in vocoder processing: Effect of number of channels and spectral region," *J. Acoust. Soc. Am.* **124**, 2272–2282.
- Traunmüller, H., and Eriksson, A. (2000). "Acoustic effects of variation in vocal effort by men, women, and children," *J. Acoust. Soc. Am.* **107**, 3438–3451.
- von Ilberg, C., Kiefer, J., Tillein, J., Pfenningdorff, T., Hartmann, R., Stürzebecher, E., and Klinke, R. (1999). "Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss," *J. Otorhinolaryngol. Relat. Spec.* **61**, 334–340.
- Whitmal, N. A., Poissant, S. F., Freyman, R. L., and Helfer, K. S. (2007). "Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience," *J. Acoust. Soc. Am.* **122**, 2376–2388.
- Whitmal, N. A., and Poissant, S. F. (2009). "The role of early reflections in the perception of spectrally-degraded speech," presented at the *Conference on Implantable Auditory Prostheses*, Lake Tahoe, CA.
- Wilson, B. S., and Dorman, M. F. (2008). "Cochlear implants: A remarkable past and a brilliant future," *Hear. Res.* **242**, 3–21.
- Zhang, T., Dorman, M. F., and Spahr, A. J. (2010). "Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation," *Ear Hear.* **31**, 63–69.