Band importance functions of listeners with cochlear implants using clinical maps

Adam K. Bosen^{a)} and Monita Chatterjee

Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131, USA

(Received 31 March 2016; revised 9 September 2016; accepted 25 October 2016; published online 16 November 2016)

Band importance functions estimate the relative contribution of individual acoustic frequency bands to speech intelligibility. Previous studies of band importance in listeners with cochlear implants have used experimental maps and direct stimulation. Here, band importance was estimated for clinical maps with acoustic stimulation. Listeners with cochlear implants had band importance functions that relied more heavily on lower frequencies and showed less cross-listener consistency than in listeners with normal hearing. The intersubject variability observed here indicates that averaging band importance functions across listeners with cochlear implants, as has been done in previous studies, may not be meaningful. Additionally, band importance functions of listeners with normal hearing for vocoded speech that either did or did not simulate spread of excitation were not different from one another, suggesting that additional factors beyond spread of excitation are necessary to account for changes in band importance in listeners with cochlear implants.

© 2016 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4967298>]

[MSS]

Pages: 3718–3727

I. INTRODUCTION

The contribution of acoustic information to speech intelligibility varies across spectral frequency. In listeners with normal hearing, the greatest contribution is provided by frequencies spanning roughly 1–2 kHz, but it is unknown whether listeners with cochlear implants also rely most on frequencies in this range for speech perception. The importance of individual frequency regions for speech perception can be measured by dividing the spectrum into discrete bands and, for each band, measuring speech intelligibility with and without that band present. The improvement in speech intelligibility gained by including a given spectral band in a speech signal indicates how important that band is for speech perception, and can be estimated across the entire spectrum to produce a band importance function (BIF). Methods such as the Articulation Index and Speech Intelligibility Index (ANSI, 1969, 1997) progressively high-pass and low-pass filter speech cues to estimate BIFs. However, there is substantial synergistic interaction between different speech bands (Warren *et al.*, 2005, 2011), which could systematically bias band importance if the co-available bands are not varied. Several alternative methods have been devised to address this issue. One approach is to simply randomize the combination of bands presented with or without the target band (Apoux and Healy, 2012; Healy *et al.*, 2013), to average band importance for the target band across all potential interactions with other bands. Others have removed individual bands from the spectrum, to observe the loss in intelligibility due to the missing information (Henry *et al.*, 2000; Sladen and Ricketts, 2015). Finally, others have introduced random levels of background noise to each band instead of removing bands, and correlated speech

intelligibility with the level of noise in each band to determine which bands are most sensitive to noise (Doherty and Turner, 1996; Mehr *et al.*, 2001; Calandruccio and Doherty, 2007). These approaches estimate the importance of each band as the combination of its independent contribution and its average synergy with all other bands, rather than just the synergy with higher or lower frequency bands as in the Articulation Index and Speech Intelligibility Index. To date, little emphasis has been placed on measuring BIFs for individual cochlear implant listeners using their clinical maps, which we address here.

Average BIFs in large populations of listeners with normal hearing are consistent across subjects (Healy *et al.*, 2013). However, listeners with cochlear implants are unlikely to demonstrate similar consistency. In cochlear implants, each channel in the device maps an acoustic band to a corresponding electrical stimulation region in the cochlea, so the resultant percept can be influenced by properties of each channel. Many properties vary both across listeners and across each channel in a listener's implant, including loudness growth (Bierer and Nye, 2014; Chatterjee *et al.*, 2000; Fu, 2005), perceptual threshold (Bierer and Nye, 2014; Pfingst *et al.*, 2004), amplitude modulation sensitivity (Chatterjee and Oberzut, 2011; Garadat *et al.*, 2012, 2013; Shannon, 1992; Zhou and Pfingst, 2014), and spread of excitation (Bierer and Faulkner, 2010; Cohen *et al.*, 2003; Hughes and Stille, 2008; Padilla and Landsberger, 2016). These differences suggest that the amount of usable information transmitted by each channel in an implant may vary in a manner that necessitates changes in the listener's reliance on the corresponding frequency bands, thereby altering a listener's BIF.

Previous studies have demonstrated that selectively modifying cochlear implant maps based on psychophysical properties of each electrode can improve speech intelligibility.

^{a)}Electronic mail: adam.bosen@boystown.org

Specifically, removing electrodes that have poor amplitude modulation detection thresholds (Garadat *et al.*, 2012, 2013) or are indiscriminable from other electrodes (Zwolan *et al.*, 1997) from a map can improve speech intelligibility. Here, we propose the first step in reversing this approach. By first directly measuring BIFs for listeners with cochlear implants we can identify channels with poor band importance and, in future work, determine the channel properties that limit band importance and design appropriate interventions to improve speech intelligibility.

The approach described here differs from previous studies of BIFs in listeners with cochlear implants by emphasizing the importance of each band within a subject’s clinical map. BIFs have been estimated with direct stimulation and experimental maps that combine several bands (Henry *et al.*, 2000; Sladen and Ricketts, 2015), and results are sometimes averaged across listeners (Sladen and Ricketts, 2015), which makes it difficult to determine the benefit each listener gains from each channel within their map. In this study, we measured BIFs of listeners with cochlear implants by presenting acoustic stimuli filtered into bands that matched each subject’s clinical map. By incorporating individual map bands and processor filtering of acoustic stimuli (as opposed to direct stimulation), our measurements estimate BIFs that more closely reflect each listener’s everyday listening conditions.

To verify the validity of this method, we measured BIFs in listeners with normal hearing using an idealized (e.g., rectangular analysis and synthesis bands) vocoder model and demonstrated strong agreement across subjects and with previous studies. We compared BIFs of listeners with normal hearing to BIFs of listeners with cochlear implants to determine how much they differed from one another. Additionally, we measured BIFs of listeners with normal hearing using a vocoder that simulates spread of excitation in monopolar stimulation in a cochlear implant to determine what effect spread of excitation would be expected to have on band importance. The work presented here demonstrates our ability to measure BIFs for cochlear implant listeners with the proposed method, which we will use in future experiments to select electrodes for further testing.

II. METHODS

This approach described here is a mix of the compound method of measuring BIFs described by Apoux and Healy (2012) and the correlational method described by Doherty and Turner (1996). Apoux and Healy (2012) chose target bands and presented pairs of trials that had the target band present or absent, in addition to a random subset of other bands. By taking the difference in speech perception when the target band was present or absent, the improvement in speech intelligibility provided by that band was estimated. A disadvantage of their approach is that it requires a large number of trials (56 trials per band) to estimate the importance of every band within one subject, so they had each listener with normal hearing perform the task for a subset of bands and averaged results across the population. The core

premise of this work is that BIFs are likely different for each cochlear implant, so averaging across listeners would render these average BIFs inadequate for predicting individual differences in speech recognition. In comparison, Doherty and Turner (1996) always had every band present, but randomly applied different levels of noise to each band. They then correlated overall speech intelligibility with the different noise levels on each band, with steeper correlations indicating a greater impact of noise and therefore a greater importance of that band for speech intelligibility. Because background noise sensitivity is variable and generally poor in listeners with cochlear implants (Friesen *et al.*, 2001), we opted to not introduce noise, but instead consider the correlation of speech intelligibility with the presence or absence of each band. We created a trial set that always had a fixed number of randomly selected bands present, with the other bands absent. Each individual band occurred an equal number of times within the trial set and the co-occurrence of pairs of bands balanced as much as feasible. We then performed a binomial regression analysis of speech intelligibility against each band’s presence or absence, to estimate the main effect each band had on speech intelligibility. This approach allowed us to collect enough data for each subject that we could estimate their individual BIF.

A. Listeners

Ten postlingually deaf listeners with cochlear implants (CI; 4 males, 6 females) and 10 listeners with normal hearing (NH; 1 male, 9 females) participated in this study. One recruited listener with a cochlear implant was replaced after poor baseline performance (17% word recognition). Demographic information for CI listeners is provided in Table I, with subjects ordered as in Fig. 2. Normal hearing listeners ranged in age from 22 to 60 (median age 26), and had pure tone thresholds <20 dB hearing level (HL) at 0.5, 1, 2, 4, and 8 kHz. All listeners were compensated hourly for participation. The study was approved by the Boys Town National Research Hospital Institutional Review Board.

TABLE I. Subject information.

Subject	Gender	Age (years)	Age at onset of deafness (years)	Duration of CI use (years)	Device
N17	F	74	8	1	CP910/CI422
N05	F	55	49	5	CP810/CI512
N15	M	61	30	2	CP810/CI422
N07	F	61	19	10	CP810/CI24R(CS)
N11 (Right ear)	M	52	42	10	Freedom/CI24RE
N11 (Left ear)			42	7	Freedom/CI24RE
C09	F	73	45	2	Naída Q70/HR90K
C06	F	75	38	20	Harmony/CI
C03 (Left ear)	M	68	26	12	Naída Q70/CII
C03 (Right ear)			26	2	Naída Q70/HR90K
C05	F	70	63	7	Harmony/HR90K
CH4	M	20	7	13	Naída Q70/CII

B. Speech material and filtering

Target stimuli were IEEE sentences (IEEE, 1969) produced by a male and a female talker. Sentences were filtered on a trial-by-trial basis to include only a subset of frequency bands, to test the importance of each band for speech intelligibility. For listeners with CIs, the range of each band was selected to match each listener's clinical map. For listeners with NH, band ranges were selected to match the lowest 20 bands of the Speech Intelligibility Index (SII) (ANSI, 1997), for comparison with previous studies of band importance. Band limits for each subject are given in Table II, and details of the filtering procedure are shown in Fig. 1.

Filtering was performed in Matlab (Mathworks, Natick, MA), as shown in Fig. 1(A). Filtering started with the generation of analysis bands by passing a sentence through a set of bandpass finite impulse response (FIR) filters, one filter for each band, with an order of 2000 and high and low cut-offs equal to the corresponding band limits (via the `filtfilt` command in Matlab, effectively producing a filter order of 4000 with no phase distortion). Filter order was chosen to minimize the introduction of spectral ripples (at lower orders) and resonant modes (at higher orders) at transition bands. For a 100 Hz cutoff the filter slope in the transition region was roughly 300 dB/octave, with a transition bandwidth of about 10–15 Hz. Filter slopes increased to about 20 000 dB/octave at an 8 kHz cutoff, with a transition bandwidth of about 10–15 Hz. The response of the analysis filters created using the SII bands to 60 s of white noise is shown in the left panel of Fig. 1(B).

For experimental trials, a fixed number of bands were pseudorandomly selected for inclusion in each trial, and the rest were removed from the stimulus. The exact number of bands selected and the procedure for selecting bands is described in Sec. II C. Additionally, baseline trials were

generated as a control condition for both NH and CI listeners. In these baseline trials, sentences were filtered as described above, but all bands were selected for inclusion in the synthesized stimulus.

For CI listeners, the selected analysis bands were summed to produce the target stimulus, as shown in the CI Bandlimited panel in Fig. 1(C). This preserved the fine temporal and spatial structure within each band that is present in typical acoustic stimuli. However, the majority of this fine structure is likely lost due to subsequent processing in listeners' implants.

For NH listeners, vocoding was performed by taking the Hilbert transform of each analysis band and low-pass filtering the result with a 300 Hz cutoff fourth order Butterworth filter to obtain the analysis band's envelope. For each sentence, a single wide band noise carrier was generated (via the `randn` command in Matlab) and normalized to have zero mean, multiplied with the envelope of each analysis band, and passed through one of two types of synthesis filter. The first type of synthesis filter had a rectangular passband identical to the filter used to generate the analysis band, which produced a nearly flat passband and heavily attenuated stopband, as shown in the left panel of Fig. 1(B). Synthesized bands were normalized so the energy in each band was equal to the energy in the corresponding analysis band (via the `rms` command in Matlab), and normalized synthesis bands were summed to produce the target stimulus. An example of the output of this synthesis method is shown in the Rect. Synthesis panel of Fig. 1(C). Rectangular passband synthesis prevents any interaction between channels, which allows for band importance to be estimated without masking or interaction between bands.

The second type of synthesis filter simulated spread of excitation using the monopolar model of Mesnildrey and

TABLE II. Band limits for each listener.

Band number	NH	N17	N5, N15, N11 (right ear)	N17	N11 (left ear)	C09	C06	C03 (both ears), C05	CH04
1	100–200	188–313	188–313	100–233	188–313	250–475	350–494	250–416	250–494
2	200–300	313–438	313–438	233–366	313–438	475–646	494–697	416–494	494–697
3	300–400	438–563	438–563	366–500	438–563	646–877	697–983	494–587	697–983
4	400–510	563–688	563–688	500–633	563–688	877–1191	983–1387	587–697	983–1387
5	510–630	688–813	688–813	633–766	688–813	1191–1617	1387–1958	697–828	1387–1958
6	630–770	813–938	813–938	766–900	813–938	1617–2196	1958–2762	828–983	1958–2762
7	770–920	938–1063	938–1063	900–1033	938–1063	2196–2982	2762–3898	983–1168	2762–3898
8	920–1080	1063–1188	1063–1188	1033–1178	1063–1188	2982–4050	3898–8700	1168–1387	3898–8700
9	1080–1270	1188–1438	1188–1313	1178–1354	1188–1438	4050–8700		1387–1648	
10	1270–1480	1438–1688	1313–1563	1354–1555	1438–1688			1648–1958	
11	1480–1720	1688–1938	1563–1813	1555–1786	1688–1938			1958–2326	
12	1720–2000	1938–2188	1813–2063	1786–2052	1938–2313			2326–2762	
13	2000–2320	2188–2563	2063–2313	2052–2380	2313–2688			2762–3281	
14	2320–2700	2563–2938	2313–2688	2380–2789	2688–3188			3281–3898	
15	2700–3150	2938–3483	2688–3063	2789–3268	3188–3688			3898–4630	
16	3150–3700	3483–3938	3063–3563	3268–3829	3688–4313			4630–8700	
17	3700–4400	3938–4563	3563–4063	3829–4486	4313–5063				
18	4400–5300	4563–5313	4063–4688	4486–5256	5063–5938				
19	5300–6400	5313–6063	4688–5313	5256–6158	5938–6938				
20	6400–7700	6063–6938	5313–6063	6158–7215	6938–7938

FIG. 1. (Color online) Stimulus synthesis procedure and examples. (A) Block diagram showing how speech stimuli were filtered. Analysis filters split the stimuli into separate bands, with band limits determined by SII bands (in NH listeners) or by individual clinical map (in CI listeners). For baseline trials, all bands were selected for inclusion in the trial, while for experimental trials a fixed number of bands were pseudorandomly selected for inclusion. For CI listeners, the output of band selection was summed to produce the final stimulus, while for NH listeners the envelope of each band was multiplied by a noise carrier filtered through either a rectangular or a monopolar model synthesis filter, then summed to produce the final stimulus. (B) The spectrum of 60 s of white noise after being passed through each filter, with filter limits determined by the SII. (C) The unfiltered spectrogram shows an example sentence from the IEEE database (“The birch canoe slid on the smooth planks,” spoken by a female talker). The CI bandlimited spectrogram shows the spectral content remaining after the example sentence was filtered for an experimental trial for a CI listener (N05), and used seven bands with limits that were unique to that listener’s clinical map. The rectangular synthesis spectrogram shows the target stimulus produced by filtering the example sentence for an experimental trial with rectangular passband synthesis (using channels 1, 5, 8, 12, 15, 18, and 20, from the SII bands). The monopolar synthesis spectrogram shows the same example, but filtered through monopolar model passband synthesis.

Macherey (2015). Briefly, this model uses models of current spread in the cochlea to guide synthesis filter design, with the objective of approximating spread of excitation from the stimulating electrode. Consequently, it produces a broad carrier which is approximately exponentially decaying from the point of stimulation (the arithmetic center of the analysis band) on a log scale, as shown in the right panel of Fig. 1(B). These filters had maximum slopes of approximately -40 dB/octave within an octave of the center frequency, which decreased at frequencies farther from the center. Unlike rectangular passband vocoding, this spread of excitation model generates interactions between different channels that may alter band importance. Spread of excitation in CI listeners is believed to alter speech perception, because incorporating spread of excitation into vocoder models produces intelligibility that is more similar to that observed in CI listeners (Fu and Nogaki, 2005; Oxenham and Kreft, 2014). By incorporating spread of excitation into a vocoder we can determine if any differences in band importance in

CI listeners can be attributed to spread of excitation. An example of the output of this synthesis method is shown in the Mono. Synthesis panel of Fig. 1(C).

C. Experiment sequence

The goal of this experiment was to measure the relative importance of each band to speech intelligibility, so we designed an experiment sequence that presented each band an equal number of times and minimized the variability in how often a pair of bands occurred within the same trial. This experiment sequence was generated in four steps: (1) Randomly choose a candidate trial set from all possible trials; (2) swap out candidate trials to try and balance the number of trials each band occurs in; (3) edit trials in the candidate set to equalize the number of trials each band occurs in; (4) repeat steps 1–3 100 times and choose the two candidate sets that have the least variability in how often pairs of bands occurred together. Details of each step are described below.

1. Sequence for users of cochlear corporation devices

Some listeners with Cochlear Corporation devices had clinical maps that contained only 20 bands, so high frequency bands were trimmed from maps that contained greater than 20 bands. Preliminary experimentation indicated that including seven bands out of the 20 available in target sentences produced some ability to accurately identify words, so each trial in this experiment was filtered to contain seven bands. A total of 400 trials was used, to produce 200 trials per talker, and each band was present in 70 trials per talker.

In step 1, we created the set of all possible trials that could be included in the experiment by generating all combinations of 20 bands taken seven at a time. Possible trials that had poor spectral range (highest and lowest bands less than 11 bands apart) and strong grouping of bands (standard deviation of band numbers less than 5) were removed from the set. Two hundred trials were randomly drawn without replacement (via the `datasample` command in Matlab) from the set to generate a candidate set of trials for the experiment. This initial candidate set contained an uneven number of occurrences of each band.

In step 2, to balance the number of times that each band occurred, trials in the candidate set that contained bands that occurred more often than desired were swapped with possible trials not in the candidate set that contained bands that occurred less often than desired. For example, if band 1 occurred 71 times and band 2 occurred 69 times in the candidate set, a trial in the candidate set that contained band 1 would be swapped for a possible trial not in the candidate set that contained band 2. This swapping procedure was unlikely to perfectly balance the number of occurrences of each band, so it was stopped after 20 000 swaps or until the standard deviation of the number of occurrences of each band in the candidate set was less than or equal to 2.25, whichever came first.

In step 3, individual bands in candidate trials that contained bands that occurred more often than desired were edited to instead contain a band that occurred less often than desired. For example if band 1 still occurred 71 times and band 2 occurred 69 times in the candidate set, a random trial that contained band 1 would be edited to replace band 1 with band 2 (as long as it did not already contain band 2). This step was performed until the candidate set contained an equal number of presentations of each band, and ignored the strong grouping limitations described in step 1.

In addition to balancing the number of times each band was presented, we also wanted to balance the number of times each pair of bands occurred together. It was not feasible to build a trial set that completely satisfied this goal, so we instead sought to minimize the range of number of times each pair of bands occurred together. In step 4, the above candidate set generation procedure was performed 100 times, and the two candidate sets that had the minimum range of number of times each pair of bands occurred together were selected as the trial sets for the experiment, for a total of 400 trials in the main experiment. One selected set of trials was used to filter sentences uttered by the male

talker using the filtering scheme described above, and the other set of trials was used to filter sentences uttered by the female talker. This approach ensured that each band was tested an equal number of times for each talker.

Stimuli were filtered prior to experimentation to contain the desired bands, as described in Sec. II B, and the resulting waveforms were saved to subject-specific wav files. After filtering, trials from the male and female talker sets were randomly interleaved to produce the final experimental trial sequence. All listeners with Cochlear devices and normal hearing performed the same trial sequence, with the only difference across listeners being the ranges for each band, which were set based on individual clinical maps or the Speech Intelligibility Index, respectively. To familiarize subjects with the speech materials and measure baseline speech intelligibility, 40 baseline trials, which included all 20 bands, were conducted before the main experiment, for each ear tested.

2. Sequence for users of advanced bionics devices

Unlike Cochlear Corporation devices, Advanced Bionics devices had fewer independent bands (two had 8, one had 9, two had 16). For listeners with these devices, we generated all combinations of 8 bands taken four at a time, 9 bands taken four at a time, or 16 bands taken six at a time. For 8 bands, this produced fewer unique combinations of bands (70 unique combinations) than the number of trials desired, so the set of all possible combinations was repeated once to produce 140 trials per talker, 280 total trials, and each channel occurred in 70 trials per talker. Similarly, for 9 bands, we used 162 trials per talker (126 unique combinations, of which 36 were selected as described for Cochlear Corporation device users to be repeated twice), 324 total trials, and each channel was present in 72 trials per talker. For 16 bands, there were more unique combinations of bands than the number of trials desired, so 200 trials per talker were selected as described for Cochlear Corporation device users, to produce 400 total trials and each channel was present in 75 trials per talker. Listeners with the same number of bands performed the same trial sequence.

D. Experiment procedure

Listeners sat in an echo-attenuated sound booth facing a loudspeaker, PC monitor, and keyboard. Speech-shaped noise was generated by concatenating the first 20 sentences in the IEEE sentence list and randomizing the phase of each frequency in the spectrum, and was used to set speaker gain such that the speech shaped noise was 65 dB sound pressure level (SPL) at the ear being tested. Listeners with normal hearing were always tested in their right ear, and wore an earplug in their left ear. Listeners with cochlear implants were first tested in their better ear (determined by clinical speech recognition scores), with the processor in their other ear removed for bilaterally implanted listeners or an earplug in their other ear for unilaterally implanted listeners. Bilaterally implanted listeners (N11 and C03) performed a smaller second study (240 trials instead of 400) with their worse ear, to compare results within the same listener.

The experiment was conducted by a custom Matlab (Mathworks, Natick, MA) script running on the PC attached to the monitor and keyboard in the sound booth. Trials were initiated by a button pressed by the listener. A filtered sentence was played once, listeners typed any words they heard in the sentence, and pressed a second button to enter their response. Listeners were not permitted to repeat sentences, and no feedback was provided. Trials were presented in blocks of 80, and listeners were allowed to proceed at their own pace. Responses were manually scored offline based on the number of correct keywords out of the five keywords in each sentence. Responses that contained insubstantial errors (i.e., typos, “reaad” instead of “read” and homophones, “piec” instead of “peace”) were counted as correct.

E. Analysis

Binomial logistic regression [model: $\text{glm}(\text{words correct}/\text{total words} \sim \text{Band 1 present} + \dots + \text{Band n present})$, binomial family, logit link] performed in R (www.r-project.org) was used to estimate the contribution of each band to speech intelligibility. This regression produced estimates of the improvement in log odds of correctly identifying words that each band contributed, which we interpret as the relative importance of that band for speech perception. Only the main effect of each band (independent contribution of each band and its average interaction with all other bands) is captured by this analysis. Due to the relatively large number of

bands, band interactions could not be reliably estimated with the available data.

III. RESULTS

Figures 2 and 3 show the BIFs obtained from listeners with cochlear implants and normal hearing, respectively. Each horizontal bar represents the band importance (in log odds) and analysis filter bandwidth for each band, and each vertical line represents the 95% confidence intervals for band importance. Log odds greater than zero indicate that a band contributed to speech intelligibility, whereas values less than zero indicate that a band did not aid speech intelligibility. Subjects within each group are ordered from worst to best performance on the baseline task, which is reported as the percentage of words correct in each panel. Performance in the experimental condition is reported in each panel as well. Results from listeners with cochlear implants that performed the task for both ears are grouped together and ordered relative to other listeners by their baseline performance in their better ear.

Listeners with cochlear implants had BIFs that differed across listeners and across ears within the same listener (Fig. 2). As shown, BIFs varied substantially across listeners, with no evident trend in BIF shape. As expected based on previous studies in listeners with normal hearing (Healy *et al.*, 2013), most listeners with cochlear implants had channels with high importance (i.e., high log odds) in the

FIG. 2. Band importance functions of listeners with cochlear implants. Cochlear Corporation device users are grouped in the left panel, and Advanced Bionics device users in the right panel.

FIG. 3. Band importance functions of listeners with normal hearing listening to vocoded speech synthesized with rectangular or monopolar model passband synthesis filters.

1–2 kHz region. However, many also appear to rely heavily on lower frequency channels, in the region of 200–400 Hz.

Caution should be used in interpreting negative log odds values. Negative values could be caused by a channel hindering speech perception, or, because there was always a fixed number of bands present on each trial, negative values could be due to that channel taking the place of a channel that would have provided more speech information. At most, negative and zero log odds indicate that a channel does not aid speech comprehension, but cannot be used to determine if a channel is hindering speech comprehension in this paradigm.

No effects of talker gender on overall intelligibility were observed for listeners with cochlear implants. Talker intelligibilities in the baseline condition had an average (and range) of 72% (44%–93%) for the male talker and 75% (47%–97%) for the female talker, which were not significantly different from one another (Wilcoxon signed rank test, $V = 26$, $p = 0.33$). In the experimental condition talker intelligibility dropped to 28% (16%–53%) for the male talker and 29% (14%–63%) for the female talker, which were also not significantly different from one another ($V = 31$, $p = 0.89$).

To validate these results, we retested four listeners with cochlear implants and two listeners with normal hearing with new experimental sequences and found good agreement

between test and retest band importance functions (Supplemental Fig. 1). Additionally, a power analysis was also performed to determine how number of trials affects estimated band importance error, which demonstrated that a sufficient number of trials were used in this experiment (Supplemental Fig. 2). We also performed a sensitivity analysis to determine how experimental performance affects estimated band importance error, which demonstrated that the range of performance observed here was sufficient to produce good estimates of band importance (Supplemental Fig. 3). Details of these analyses are described in the supplemental materials.¹

In comparison, the band importance functions obtained from listeners with normal hearing (Fig. 3) followed the characteristic inverted “U” shape that has been previously observed (ANSI, 1969, 1997; Apoux and Healy, 2012; Healy *et al.*, 2013). Band importance peaked at about 1–2 kHz, with bands outside of this frequency region contributing less to speech intelligibility. Additionally, large variations across adjacent channels within this trend are apparent and consistent across subjects, in agreement with the compound method results of Healy *et al.* (2013). The true band importance functions for each listener were assumed to be identical, and results from each subject were averaged together for each synthesis filter type. The across-subject mean and 95% confidence intervals for each band

are given as the last panels in Fig. 3. Good agreement across subjects indicate that the method used here reliably measures band importance with sufficient precision to allow for estimating subject specific BIFs within a single experimental session. Results did not appear to differ substantially across synthesis filter types. Unlike listeners with cochlear implants, listeners with normal hearing did not show high band importance for lower frequency bands.

Overall performance was not sensitive to synthesis filter type. Speech intelligibility in the baseline condition had an average (and range) of 97% (97%–98%) for rectangular synthesis and 96% (94%–98%) for monopolar model synthesis. No statistical analysis is given because these results are at ceiling. In the experimental condition speech intelligibility dropped to 51% (42%–59%) for rectangular synthesis and 56% (46%–64%) for monopolar model synthesis, which were not different from one another (Wilcoxon rank sum test, $W = 6.5$, $p = 0.25$).

In contrast with listeners with cochlear implants, listeners with normal hearing were slightly better at understanding the male talker in the experimental condition. Talker intelligibilities in the baseline condition had an average (and range) of 98% (94%–99%) for the male talker and 96% (93%–98%) for the female talker across both synthesis filter types. Again, no statistical analysis is given because these results are at ceiling. In the experimental condition talker intelligibility dropped to 58% (45%–65%) for the male talker and 49% (38%–63%) for the female talker, which were significantly different from one another (Wilcoxon signed rank test, $V = 55$, $p = 0.006$, threshold = 0.013 after Bonferroni correction). However, the validity of this statistic will need to be confirmed with a larger sample size before any conclusions can be drawn.

IV. DISCUSSION

Listeners with normal hearing had similar BIFs to one another, indicating that the bands that are important for speech perception are the same across these listeners. In comparison, listeners with cochlear implants showed high variability in their BIFs, even when obtained from the same listener in two different ears. This strong cross-listener variability indicates that averaging BIFs across cochlear implant listeners would miss features of individual band importance functions. That being said, it appears that most listeners with cochlear implants rely heavily on the lowest frequency bands, a trend not observed in listeners with normal hearing. This observation suggests that cochlear implant BIFs may not be explainable simply as normal hearing BIFs degraded by channel properties, but rather reflect a distinct strategy for speech perception. Cross-ear differences within listeners indicate that band importance is at least partially dependent on the specific properties of the tested ear, and cannot be solely accounted for by a common mechanism across ears.

Previous work on BIFs in listeners with cochlear implants has removed individual frequency bands from an experimental map, creating a hole in the speech spectrum (Henry *et al.*, 2000; Sladen and Ricketts, 2015). These experiments used five channel experimental maps and CNC

words (Henry *et al.*, 2000) and six channel experimental maps with nonsense syllables (Sladen and Ricketts, 2015), and found band importance functions that varied little across bands or listeners. The hole method is essentially a variation of the method demonstrated here, where $n-1$ of n bands are chosen on each trial. However, we contend that it is difficult to estimate the true importance of the dropped band when it is the only information removed, since the redundancy present in speech (Warren *et al.*, 2005, 2011) is likely to compensate for the dropped band, and thereby underestimate that band's importance for speech. By removing several bands at once, we provide a greater challenge by removing much of this redundancy, and are less likely to underestimate each band's importance. Another experiment used different levels of noise applied within each band of listeners' clinical maps with six channel processors, and correlated speech intelligibility with the noise level in each band to determine band importance (Mehr *et al.*, 2001), which similarly challenges the redundancy in speech. Results of this experiment demonstrated heterogeneity in band importance across listeners with cochlear implants similar to the results observed here. They further validated their results by using the correlational method results to select channels for removal in the hole method, and demonstrated good agreement across the two methods. However, the speech materials they used were individual consonant-vowel syllables, which may not be directly comparable to our IEEE sentence based results.

By measuring BIFs of clinical maps we ensure that subjects have had extensive experience listening to the tested map, and therefore have had time to fully adapt to listening with it. Testing subjects with maps they have fully adapted to is a likely explanation for the high variability in band importance across subjects, since they have had time to learn to use the cues transmitted through their individual implants and auditory systems. Given the relatively small variation in band limits across listeners (Table II), we believe it is unlikely that the band limits can account for much of the observed variation, although additional experimentation would be needed to verify this claim.

Most listeners with cochlear implants had high band importance for the lowest channels in their map. This is surprising, because listeners with normal hearing did not show much importance for low frequency channels. This increased reliance on low frequency channels may be due to the fact that cochlear implants have relatively poor ability to convey place of excitation cues, but are better at conveying temporal envelope cues (Xu and Pfingst, 2008). Experiments in listeners with normal hearing with transposed tones, sounds where low rate modulation is applied to a high frequency carrier, demonstrated that correct (i.e., low frequency) place of excitation is necessary to perceive the fundamental frequency of low rate modulation (Oxenham *et al.*, 2004). This type of mechanism in listeners with cochlear implants would make temporal envelope cues more valuable in lower frequency channels, and thereby make those low frequency channels more important for speech perception, particularly if higher frequency spectral cues were unavailable. Alternatively, this reliance on low frequency channels could reflect the importance of the place of stimulation, rather than the acoustic

frequencies mapped to that place. Shallow implant insertion depth would misalign band frequencies and place of excitation within the cochlea, which would cause low frequency analysis bands to be mapped to a stimulation site more basal (i.e., to a higher frequency place) than such stimulation would excite in normal hearing. All of the listeners with cochlear implants were post-lingually deafened, so it is possible that they developed speech sensitivity that is location specific (and would be maximal at the place that normally corresponds to 1–2 kHz), and that the lower stimulus frequencies (200–400 Hz) are being presented to a more sensitive region than they would be in normal hearing due to shallow insertion.

The current experiment design tried to match the fraction of active channels across listeners in the experimental condition, but this produced a broad range of performance across listeners with cochlear implants. Ideally, performance would be uniformly distributed from 0 to 100% word recognition within each listener. While this goal may be infeasible by only manipulating the fraction of active channels, adjusting this parameter such that performance is more uniformly distributed for each subject would improve the accuracy of estimated BIFs. However, there is no guarantee that manipulating the fraction of active channels will not alter the shape of the listener's BIF, which would need to be investigated in future studies.

Surprisingly, there was no difference in BIFs obtained from listeners with normal hearing between the rectangular passband and monopolar model passband synthesis filter vocoders. It is likely that the monopolar model synthesis used here did not produce sufficient spread of excitation to necessitate alternate listening strategies. Other studies that have demonstrated a deterioration of speech intelligibility due to spread of excitation (Fu and Nogaki, 2005; Oxenham and Kreft, 2014) have used synthesis filter slopes shallower (fixed -6 or -24 dB/octave and -8 dB/octave, respectively) than the monopolar model synthesis in this experiment (which had a maximum slope of approximately -40 dB/octave within an octave of the center frequency). Alternatively, it is possible that listeners did not have enough experience to adapt their BIFs to compensate for spectral smearing. However, this seems unlikely, since performance did not differ across synthesis filter types.

The BIFs presented here are not intended to generalize to other types of stimuli or listening environments, but rather reflect the specific conditions used in this experiment. Since BIFs of listeners with cochlear implants were obtained with matched stimuli results are comparable across the two groups, but these results are likely sensitive to listening environment and talker. More experimental conditions must be tested before the real-world relevance of the BIFs measured here can be determined. Additionally, future psychophysical and physiological work will focus on identifying and distinguishing between the underlying factors that determine band importance and where they occur in the auditory pathway.

ACKNOWLEDGMENTS

This work was supported by NIH Grant Nos. R01 DC014233 and T32 DC000013. Subject recruitment was

facilitated by the Human Subjects Research Core of NIH Grant No. P30-DC004662. Discussions with Jeffrey L. Simmons and Ryan W. McCreery guided development of this project, Aditya M. Kulkarni and Sara A. Damm assisted in data collection, and Tim Vallier assisted in the creation of the github repository where the code developed here is publicly available. We are grateful to all listeners who participated in this study. We thank the A.E. and two anonymous reviewers for their thoughtful comments on a previous version of this manuscript.

¹See supplementary material at <http://dx.doi.org/10.1121/1.4967298>.

- ANSI (1969). *American National Standard Methods for the Calculation of the Articulation Index* (American National Standards Institute, New York).
- ANSI (1997). *American National Standard: Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).
- Apoux, F., and Healy, E. W. (2012). "Use of a compound approach to derive auditory-filter-wide frequency-importance functions for vowels and consonants," *J. Acoust. Soc. Am.* **132**(2), 1078–1087.
- Bierer, J. A., and Faulkner, K. F. (2010). "Identifying cochlear implant channels with poor electrode-neuron interface: Partial tripolar, single-channel thresholds and psychophysical tuning curves," *Ear Hear.* **31**(2), 247–258.
- Bierer, J. A., and Nye, A. D. (2014). "Comparisons between detection threshold and loudness perception for individual cochlear implant channels," *Ear Hear.* **35**(6), 641–651.
- Calandruccio, L., and Doherty, K. A. (2007). "Spectral weighting strategies for sentences measured by a correlational method," *J. Acoust. Soc. Am.* **121**(6), 3827–3836.
- Chatterjee, M., Fu, Q. J., and Shannon, R. V. (2000). "Effects of phase duration and electrode separation on loudness growth in cochlear implant listeners," *J. Acoust. Soc. Am.* **107**(3), 1637–1644.
- Chatterjee, M., and Oberzutz, C. (2011). "Detection and rate discrimination of amplitude modulation in electrical hearing," *J. Acoust. Soc. Am.* **130**(3), 1567–1580.
- Cohen, L. T., Richardson, L. M., Saunders, E., and Cowan, R. S. C. (2003). "Spatial spread of neural excitation in cochlear implant recipients: Comparison of improved ECAP method and psychophysical forward masking," *Hear. Res.* **179**(1–2), 72–87.
- Doherty, K. A., and Turner, C. W. (1996). "Use of a correlational method to estimate a listener's weighting function for speech," *J. Acoust. Soc. Am.* **100**(6), 3769–3773.
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**(2), 1150–1163.
- Fu, Q. J. (2005). "Loudness growth in cochlear implants: Effect of stimulation rate and electrode configuration," *Hear. Res.* **202**, 55–62.
- Fu, Q. J., and Nogaki, G. (2005). "Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing," *J. Assoc. Res. Otolaryngol.* **6**(1), 19–27.
- Garadat, S. N., Zwolan, T. A., and Pfingst, B. E. (2012). "Across-site patterns of modulation detection: Relation to speech recognition," *J. Acoust. Soc. Am.* **131**(5), 4030–4041.
- Garadat, S. N., Zwolan, T. A., and Pfingst, B. E. (2013). "Using temporal modulation sensitivity to select stimulation sites for processor maps in cochlear implant listeners," *Audiol. Neurotol.* **18**(4), 247–260.
- Healy, E. W., Yoho, S. E., and Apoux, F. (2013). "Band importance for sentences and words reexamined," *J. Acoust. Soc. Am.* **133**(1), 463–473.
- Henry, B. A., McKay, C. M., McDermott, H. J., and Clark, G. M. (2000). "The relationship between speech perception and electrode discrimination in cochlear implantees," *J. Acoust. Soc. Am.* **108**(3 Pt 1), 1269–1280.
- Hughes, M. L., and Stille, L. J. (2008). "Psychophysical versus physiological spatial forward masking and the relation to speech perception in cochlear implants," *Ear Hear.* **29**(3), 435–452.
- IEEE (1969). *IEEE Recommended Practice for Speech Quality Measurements* (Institute of Electrical and Electronic Engineers, New York).

- Mehr, M. A., Turner, C. W., and Parkinson, A. (2001). "Channel weights for speech recognition in cochlear implant users," *J. Acoust. Soc. Am.* **109**(1), 359–366.
- Mesnildrey, Q., and Macherey, O. (2015). "Simulating the dual-peak excitation pattern produced by bipolar stimulation of a cochlear implant: Effects on speech intelligibility," *Hear. Res.* **319**, 32–47.
- Oxenham, A. J., Bernstein, J. G. W., and Penagos, H. (2004). "Correct tonotopic representation is necessary for complex pitch perception," *Proc. Natl. Acad. Sci. U.S.A.* **101**(5), 1421–1425.
- Oxenham, A. J., and Kreft, H. A. (2014). "Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing," *Trends Hear.* **18**, 1–14.
- Padilla, M., and Landsberger, D. M. (2016). "Reduction in spread of excitation from current focusing at multiple cochlear locations in cochlear implant users," *Hear. Res.* **333**, 98–107.
- Pfingst, B. E., Xu, L., and Thompson, C. S. (2004). "Across-site threshold variation in cochlear implants: Relation to speech recognition," *Audiol. Neuro-Otol.* **9**(6), 341–352.
- Shannon, R. V. (1992). "Temporal modulation transfer functions in patients with cochlear implants," *J. Acoust. Soc. Am.* **91**(4 Pt 1), 2156–2164.
- Sladen, D. P., and Ricketts, T. A. (2015). "Frequency importance functions in quiet and noise for adults with cochlear implants," *Am. J. Audiol.* **24**(2), 477–486.
- Warren, R. M., Bashford, J. A., and Lenz, P. W. (2005). "Intelligibilities of 1-octave rectangular bands spanning the speech spectrum when heard separately and paired," *J. Acoust. Soc. Am.* **29**(6), 997–1003.
- Warren, R. M., Bashford, J. A., and Lenz, P. W. (2011). "An alternative to the computational Speech Intelligibility Index estimates: Direct measurement of rectangular passband intelligibilities," *J. Exp. Psychol. Hum. Percept. Perform.* **29**(6), 997–1003.
- Xu, L., and Pfingst, B. E. (2008). "Spectral and temporal cues for speech recognition: Implications for auditory prostheses," *Hear. Res.* **242**, 132–140.
- Zhou, N., and Pfingst, B. E. (2014). "Effects of site-specific level adjustments on speech recognition with cochlear implants," *Ear Hear.* **35**(1), 30–40.
- Zwolan, T. A., Collins, L. M., and Wakefield, G. H. (1997). "Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects," *J. Acoust. Soc. Am.* **102**(6), 3673–3685.