

The noise susceptibility of various speech bands

Sarah E. Yoho, Frédéric Apoux, and Eric W. Healy

Citation: [The Journal of the Acoustical Society of America](#) **143**, 2527 (2018); doi: 10.1121/1.5034172

View online: <https://doi.org/10.1121/1.5034172>

View Table of Contents: <http://asa.scitation.org/toc/jas/143/4>

Published by the [Acoustical Society of America](#)

Articles you may be interested in

[Correlations between otoacoustic emissions and performance in common psychoacoustical tasks](#)

[The Journal of the Acoustical Society of America](#) **143**, 2355 (2018); 10.1121/1.5030999

[Discovering acoustic structure of novel sounds](#)

[The Journal of the Acoustical Society of America](#) **143**, 2460 (2018); 10.1121/1.5031018

[Arrays of rectangular subcritical speech bands: Intelligibility improved by noise-vocoding and expanding to critical bandwidths](#)

[The Journal of the Acoustical Society of America](#) **143**, EL305 (2018); 10.1121/1.5034170

[Effect of audibility on better-ear glimpsing as a function of frequency in normal-hearing and hearing-impaired listeners](#)

[The Journal of the Acoustical Society of America](#) **143**, 2195 (2018); 10.1121/1.5031007

[Assessing the importance of several acoustic properties to the perception of spontaneous speech](#)

[The Journal of the Acoustical Society of America](#) **143**, 2255 (2018); 10.1121/1.5031123

[Differences in common psychoacoustical tasks by sex, menstrual cycle, and race](#)

[The Journal of the Acoustical Society of America](#) **143**, 2338 (2018); 10.1121/1.5030998

The noise susceptibility of various speech bands

Sarah E. Yoho,^{a)} Frédéric Apoux,^{b)} and Eric W. Healy

Department of Speech and Hearing Science, The Ohio State University, Columbus, Ohio 43210, USA

(Received 12 October 2017; revised 21 February 2018; accepted 6 April 2018; published online 30 April 2018)

The degrading influence of noise on various critical bands of speech was assessed. A modified version of the compound method [Apoux and Healy (2012) *J. Acoust. Soc. Am.* **132**, 1078–1087] was employed to establish this noise susceptibility for each speech band. Noise was added to the target speech band at various signal-to-noise ratios to determine the amount of noise required to reduce the contribution of that band by 50%. It was found that noise susceptibility is not equal across the speech spectrum, as is commonly assumed and incorporated into modern indexes. Instead, the signal-to-noise ratio required to equivalently impact various speech bands differed by as much as 13 dB. This noise susceptibility formed an irregular pattern across frequency, despite the use of multi-talker speech materials designed to reduce the potential influence of a particular talker's voice. But basic trends in the pattern of noise susceptibility across the spectrum emerged. Further, no systematic relationship was observed between noise susceptibility and speech band importance. It is argued here that susceptibility to noise and band importance are different phenomena, and that this distinction may be underappreciated in previous works. © 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5034172>

[SHF]

Pages: 2527–2534

I. INTRODUCTION

Speech-shaped noise is often used to mask speech because it produces the same long-term average signal-to-noise ratio (SNR) in each frequency band. The long-held assumption is that the same amount of masking is also produced in each band, because the susceptibility of speech to noise is the same across the frequency spectrum. Accordingly, as Miller (1947) stated seven decades ago, noise matching the long-term average amplitude spectrum of speech has long been considered the most effective masker of speech.

This assumption is reflected in the ANSI-Standard Speech Intelligibility Index (SII; ANSI, 1997). The SII provides an audibility factor (A), which scales the contribution of a speech band (its importance) by its audibility, or the extent to which the band is available in the presence of background noise. The contribution of a speech band is typically considered to be entirely intact when its SNR is at or above 18 dB, and the contribution is typically considered to be entirely absent when the SNR is at or below –12 dB. Each increase in SNR of 3 dB typically corresponds to an increase in A of 0.1. Important for the purposes of the current study, this audibility factor is constant across all bands.

On one hand, there is reason to believe this assumption that the susceptibility of speech to noise is the same across the spectrum. The psychoacoustic sensitivity to pure-tone signals in noise maskers does not vary greatly as a function of signal frequency (Hawkins, Jr. and Stevens, 1950). French and Steinberg (1947) also describe the masking of pure tones

by bands of noise as relying solely on the SNR in the critical band and the threshold of the signal in quiet.

However, there is also reason to believe that the assumption of equal noise susceptibility of speech is not true. Speech bands in different regions of the spectrum possess dramatically different acoustic characteristics and code different types of linguistic cues in different ways. These different acoustic codings may be more or less susceptible to corruption by noise. Data also exist to suggest that the influence of noise on speech is not entirely uniform across the spectrum. The Articulation Index (ANSI, 1969) crossover frequency, which divides the spectrum into equally contributing halves, may be dependent on SNR. Webster and Klumpp (1963) concluded that the crossover frequency can decrease by as much as one octave in noise relative to the value in quiet observed by French and Steinberg (1947). Miller and Nicely (1955) observed different recognition performance across various consonants, despite the use of a constant noise and SNR. This could reflect the different spectral compositions of the various consonants. Studebaker and Sherbecoe (2002) examined relative importance at different speech intensities in fixed-level noise. The resulting intensity-importance functions were similar to that in the SII, but several differences were observed. Notably, the functions were different for different speech frequencies, suggesting a possibly different influence of noise across frequency. Finally, Apoux and Bacon (2004) found the pattern of relative importance for four vocoded speech bands to differ across quiet versus noise conditions, potentially reflecting a different impact of the noise across the four bands.

The current concept of noise susceptibility of speech has important implications for the more established concept of speech band importance. The relative contributions of various bands of speech to overall intelligibility have been studied extensively and are reflected in the band-importance functions in the SII. However, this work is limited in its ability to

^{a)}Present address: Department of Communicative Disorders & Deaf Education, Utah State University, Logan, UT 84322, USA. Electronic mail: sarah.leopold@usu.edu

^{b)}Present address: Department of Otolaryngology–Head & Neck Surgery, The Ohio State University Wexner Medical Center, Columbus, OH 43212, USA.

assess the detrimental influence of noise on various bands of speech. The standard technique for assessing band importance (Studebaker *et al.*, 1987; Studebaker and Sherbecoe, 1991) requires the use of background noise to control overall intelligibility and avoid ceiling effects during testing. In the correlational method (Doherty and Turner, 1996), speech band importance is assessed by measuring the detrimental impact of noise on various speech bands. In this method, a band is considered important if it is resistant to the corrupting influence of noise. Thus, many existing techniques to establish the relative importance of various speech bands confound band importance with noise susceptibility.

It is seemingly important to examine these two factors separately, to determine the extent to which noise susceptibility varies across the spectrum independently from band importance. A recently developed technique to derive speech band importance allows this distinction. In the compound method (Apoux and Healy, 2012; Healy *et al.*, 2013), background noise is not required, and so the resulting band-importance functions are not confounded by potentially differing influences of noise susceptibility across the spectrum. The method consists of two speech-intelligibility measurements, one in which a band of interest is presented with n other randomly distributed bands (band-present condition), and another where the n other bands are presented without the band of interest (band-absent condition). These paired conditions are repeated over many trials, with new random draws to determine the frequency positions of the n other bands. The importance of the band of interest is then determined by the relative difference between its band-present and band-absent scores.

In the current study, the compound method was adapted to examine the noise susceptibility of various speech frequencies. This was done by adding noise at various levels to each speech band in turn. When the compound method is used and no background noise is employed, the intelligibility obtained when a given target speech band is present is higher than when that band is absent. A certain amount of noise added to the target speech band will leave the band unaffected, and scores should be equal to the band-present-in-quiet score. A larger amount of noise will obliterate the influence of that target speech band, and the score should be equal to the band-absent-in-quiet score. To obtain a sensitive measure of noise susceptibility for each speech band, the amount of noise required to reduce scores half-way from band-present-in-quiet to band-absent-in-quiet was obtained currently. This noise susceptibility was then compared across speech bands. For example, if band x requires a certain amount of noise to be impacted by a certain amount, and if band y requires a larger amount of noise to be impacted by that same amount, then band x has a greater susceptibility to noise than band y . This noise sensitivity may be expressed as equivalent SNRs—those required to affect different speech bands equivalently.

II. METHOD

A. Subjects

Forty normal-hearing listeners between the ages of 19 and 33 yr (mean = 21.0) participated in this experiment.

Thirty-five were female. The subjects were recruited from courses at The Ohio State University and received either course credit or a monetary incentive for participation. All had pure-tone audiometric thresholds at or below 20 dB hearing level (HL) at octave frequencies from 250 to 8000 Hz (ANSI, 2004, 2010). None had previous exposure to the sentence materials used in this study.

B. Stimuli

The speech materials were sentences from the IEEE database (IEEE, 1969). The corpus is composed of 720 sentences, each containing five scoring key words. The original 22.05 kHz, 16-bit recordings spoken by 10 different talkers (5 male and 5 female) judged to have a general American dialect were used. The sentences were filtered into the 21 critical bands specified in the SII (see Table I). Filter orders were chosen to approximately equate filter slopes across bands in dB/oct and to ensure minimal acoustic band overlap. This was accomplished through the use of high-order fir filters, which preserve the amplitude and phase response within the passbands (see Healy, 1998). Filter orders were adjusted for each band to produce approximately equal slopes that exceeded 1000 dB/oct (see Healy *et al.*, 2013). These orders ranged from 1000 for the highest frequency band to 10 000 for the lowest frequency band. The stimuli were filtered in the forward and reverse direction so that no group delays were introduced. The relative spectrum level of each speech band was maintained. A corresponding band of Gaussian noise was created for each speech band using the same cutoffs and orders as for each speech band. The noise had a 10-ms raised cosine rise/fall and started at least 300 ms prior to each sentence to avoid possible effects of overshoot (Bacon and Liu, 2000). All processing was performed in MATLAB.

TABLE I. Band divisions for the 21 SII critical bands.

Band	Center frequency (Hz)	Band limits (Hz)
1	150	100–200
2	250	200–300
3	350	300–400
4	450	400–510
5	570	510–630
6	700	630–770
7	840	770–920
8	1000	920–1080
9	1170	1080–1270
10	1370	1270–1480
11	1600	1480–1720
12	1850	1720–2000
13	2150	2000–2320
14	2500	2320–2700
15	2900	2700–3150
16	3400	3150–3700
17	4000	3700–4400
18	4800	4400–5300
19	5800	5300–6400
20	7000	6400–7700
21	8500	7700–9500

C. Procedure

A modified version of the compound band-importance method was employed to determine noise susceptibility for each speech band. In this modification, the target speech band was always present, along with noise at different SNRs. It was always presented with four other speech bands, randomly selected from the 21 bands on each trial and for each subject. For each trial, noise was introduced to the target band to achieve one of six SNRs: -12 , -8 , -4 , 0 , 4 , or 8 dB. Conditions were blocked such that each SNR for a given target band was completed before moving on to a different target band. The order in which SNRs and target bands were heard was randomized for each subject. A total of 10 sentences was employed for each target speech band at each SNR, by using one sentence from each talker. The sentence and talker were chosen randomly without replacement for each trial.¹ Due to the very small band-present-in-quiet minus band-absent-in-quiet intelligibility difference observed for band 1 (center frequency 150 Hz), its noise susceptibility was not assessed. The subjects were randomly divided into two groups of 20 subjects each, and assigned to hear all of the even-numbered bands or all of the odd-numbered bands. A total of 60 conditions were therefore presented to each subject (10 bands \times 6 SNRs). Test duration was approximately 2 h, with most individuals completing the experiment over two sessions within a two-week period.

The stimuli were converted to analog form using a PC and Echo Gina 3 G D/A converters. They were presented diotically via Sennheiser HD 280 circumaural headphones. Broadband sentences (21 summed speech bands) were set to play back at 70 dBA at each earphone using a flat-plate coupler (Larson Davis AEC 101) and ANSI Class 1 sound level meter (Larson Davis 824). Subjects were seated in a double-walled IAC sound booth with the experimenter. They were instructed to repeat back as much of each sentence as possible to the experimenter, who recorded responses with the assistance of a custom MATLAB script. Prior to testing, subjects completed a familiarization in which they heard 20 sentences spoken by a male and female talker not heard during testing. Presented were five broadband sentences, five sentences as 11 randomly selected critical bands (no noise), and finally 10 sentences as four randomly selected critical bands. Correct/incorrect feedback was given during this familiarization stage only.

D. Baseline scores

The band-present and band-absent scores in quiet for each target speech band were drawn from [Yoho et al. \(2018\)](#), using the methods of [Apoux and Healy \(2012\)](#) and [Healy et al. \(2013\)](#). These compound-method procedures were essentially identical in every important way to the modified compound method employed currently. These baseline scores were determined as follows: Three groups of 20 normal-hearing subjects having characteristics similar to those employed currently were employed (ages 19–37 yr, mean = 21.8, 55 females, audiometric thresholds of 20 dB HL or better, no previous exposure to IEEE sentences). They were assigned to three subsets of target-band conditions (bands 1–7, bands 8–14, or bands 15–21).

Stimuli were the same 10-talker IEEE sentence recordings employed currently, filtered into the same 21 critical bands using the same filtering parameters as in the current experiment. For each target-band condition, the band of interest was presented along with four other bands having frequency positions determined randomly for each trial and subject (target band-present condition). Each of these trials was paired with a contiguously presented trial in which the same “other” bands were presented without the target band (target band-absent condition). No noise was employed. Conditions were blocked by target band and randomized. Sentence-to-condition correspondence and order of band present/band absent in each paired trial were also randomized. Each of the seven target-band conditions included two (one band-present, one band-absent) sentences from each of the 10 talkers, for a total of 20 sentences, with the order of talkers randomized within each block. The level of the broadband speech (all 21 bands) was set to 70 dBA using the apparatus employed in the current experiment. The test setting, instructions, and playback apparatus were also identical. A familiarization identical to that of the current experiment preceded data collection. For each target-band condition, a band-present and a band-absent score (in quiet) were calculated and averaged across subjects.

III. RESULTS

Figure 1 shows sentence intelligibility in percent-correct keywords as a function of target-band SNR, averaged across both subject groups and all 20 target-band conditions (average of the data displayed in Figs. 2–4). Group-mean intelligibility at each SNR is displayed as filled symbols, along with

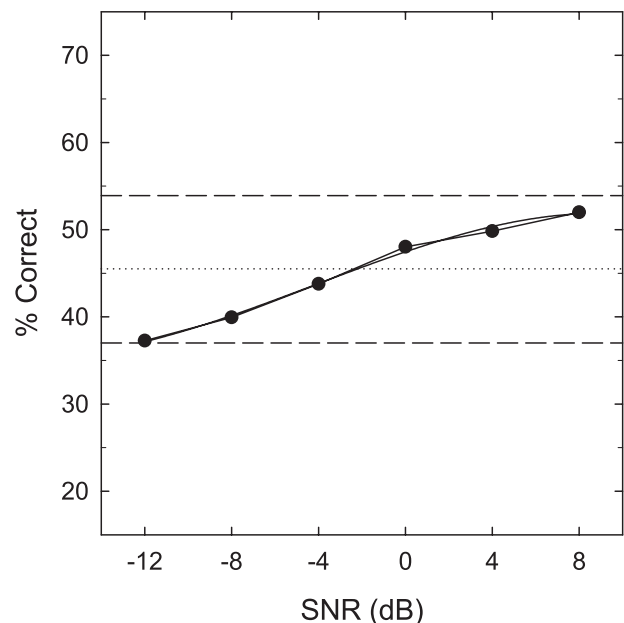


FIG. 1. Group-mean sentence intelligibility in percent correct as a function of target-band signal-to-noise ratio averaged across all 20 critical bands tested. Also shown is the third-order regression fit to these data. The top dashed line represents group-mean band-present-in-quiet sentence intelligibility across all 20 critical bands tested, and the bottom dashed line represents group-mean band-absent-in-quiet sentence intelligibility across all 20 critical bands tested (reference-line data from [Yoho et al., 2018](#)). The dotted line represents the midway point between the plotted band-present and band-absent scores.

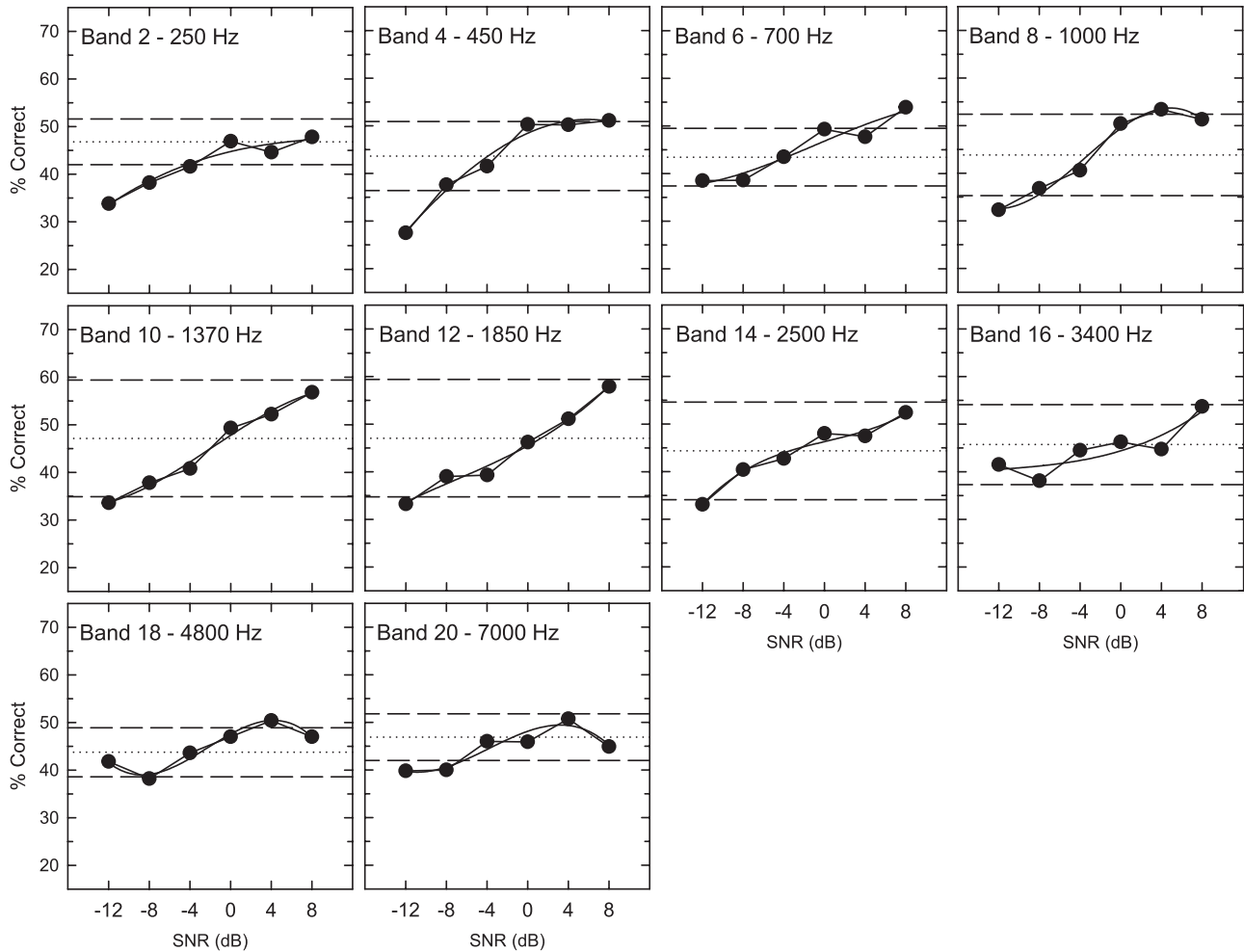


FIG. 2. As Fig. 1, except data are plotted separately in each panel for the 10 even-numbered critical bands tested. Band number and center frequency are provided.

a third-order regression line fit to these data. The top and bottom dashed lines represent band-present and band-absent scores in quiet, respectively. These scores are also averaged across all 20 target-band conditions and are from [Yoho et al. \(2018\)](#). The dotted line represents the half-way point between these band-present and band-absent scores. Because the band-present and band-absent scores straddled the linear portion of the psychometric intelligibility function near 50% correct, the arithmetic mean was used to define the half-way point. Figure 1 confirms that the SNR conditions employed currently were sufficient to observe scores that matched well on average the band-present and band-absent scores in quiet, and that a smooth function between these two end-points is observed as a function of increasing noise level.

Figures 2 and 3 show similar data, but for each target band individually. The symbols in each panel represent data from the current subjects hearing each target band at various SNRs, the curve is a third-order regression fit to these data, the dashed reference lines in each panel represent band-present and band-absent scores in quiet for that target band (from [Yoho et al., 2018](#)), and the dotted line in each panel represents the half-way point. The point at which the regression line intersected the half-way point was determined for each band, and these values are given in Table II. This

intersection reflects the SNR required to reduce intelligibility resulting from that band by half.

It is important to note that the functions displayed in Figs. 2 and 3 were derived using a group of subjects different from those used to determine the band-present- and band-absent-in-quiet reference lines. Whereas most functions involving SNRs of -12 to 8 dB span these reference lines with reasonable accuracy, the functions for the two of the most extreme bands (bands 2 and 20) matched these previous reference data with less accuracy. This is likely due to the narrow range of band-present minus band-absent scores resulting from their relatively low importance. As a result, additional within-subjects control conditions were implemented to re-assess bands 2 and 20. These conditions involved five normal-hearing subjects (ages 19–20 yr, mean = 19.4, 5 females, audiometric thresholds of 20 dB HL or better, no previous exposure to IEEE sentences) who did not take part in the other experiments. For each of these two bands, they heard SNRs of -12 to 12 dB in 4-dB steps, plus band-present and band-absent conditions in quiet. Each subject heard 30 of the 10-talker IEEE sentences (three sentences per talker) in each of these 18 conditions, for a total of 540 sentences, with the sentence-to-condition correspondence randomized for each subject. As in the main experiment, the order of SNR and target-band conditions was blocked

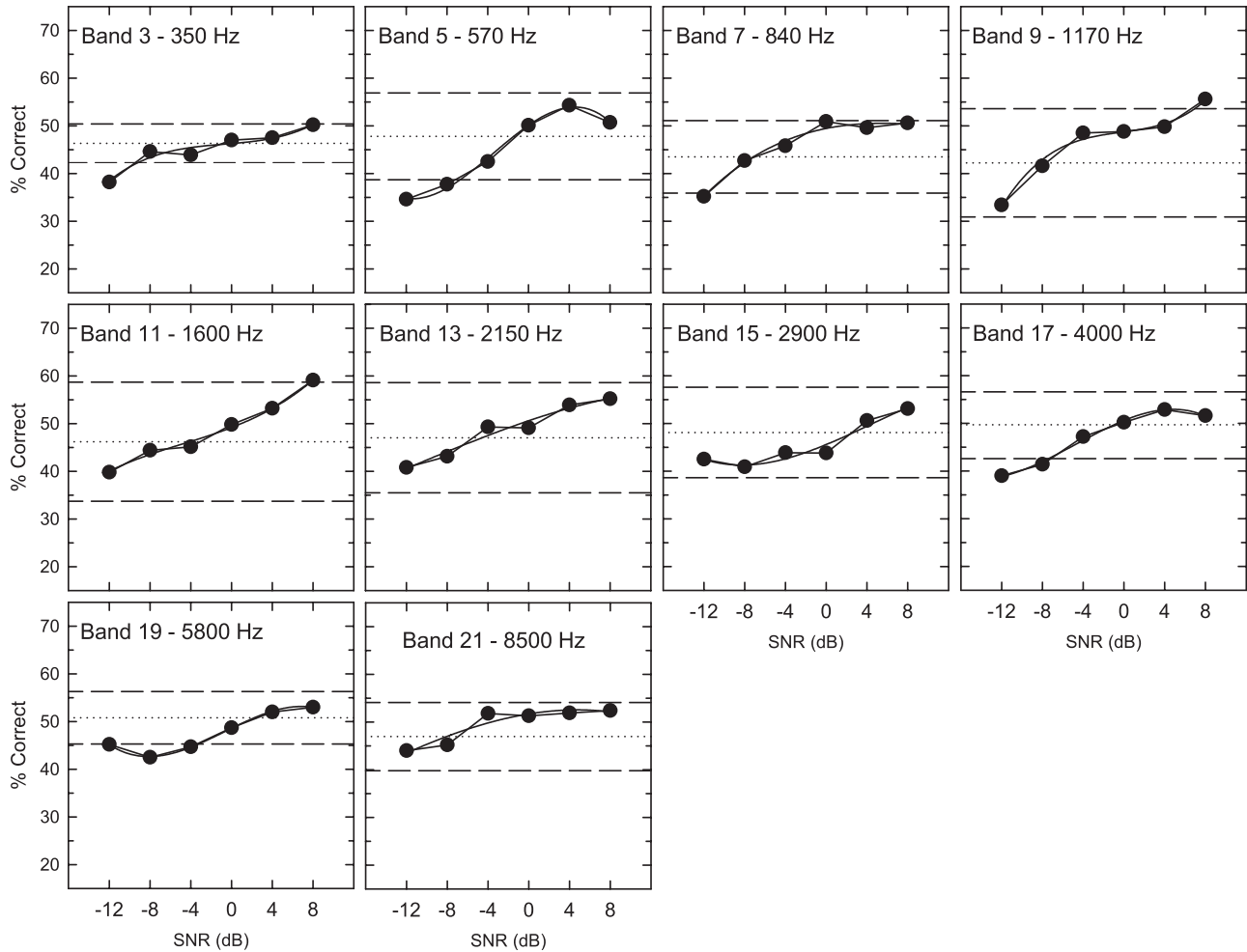


FIG. 3. As Fig. 2, but for the 10 odd-numbered critical bands tested.

and randomized for each subject. All other methods and apparatus were identical to those employed in the main experiment. Results are shown in Fig. 4. The data obtained in these conditions displayed a greater degree of agreement between the band-present/band-absent-in-quiet values and the function relating intelligibility to SNR of the target band. The resulting noise-susceptibility values were within 1.6 dB of the values originally obtained and displayed in Fig. 2. But because of the closer correspondence obtained in the control conditions, these susceptibility values were used to best represent bands 2 and 20. Noise-susceptibility values for each critical speech band are located in Table II and Fig. 5 as shaded columns.

Also shown in Fig. 5 are the noise-susceptibility data smoothed using a three-band rectangular window. Each of 20 smoothed noise-susceptibility values was obtained by averaging each target-band value with that of the two adjacent bands. The lowest and highest frequency bands were averaged with the one adjacent band. These windowed values were then fit using a spline curve.

IV. DISCUSSION

The present data demonstrate the noise susceptibility of various bands of speech. The degree of vulnerability, or conversely the degree of robustness, to the detrimental influence

of extraneous noise was systematically evaluated for 20 critical bands spanning the speech spectrum. To examine this, an adaptation of the compound method (Apoux and Healy, 2012; Healy *et al.*, 2013) was employed. Sentence intelligibility was measured while the band of interest was presented along with four other bands randomly distributed in frequency from trial-to-trial.

New to the current manipulation is the addition of noise to the target band at different SNRs. Intelligibility as a function of SNR within the target band forms a psychometric function, which asymptotes at the band-present-in-quiet score at the top of the function, and at the band-absent-in-quiet score at the bottom of the function. This is because at some favorable SNR, the target speech band is essentially unaffected by noise, and at some unfavorable SNR, the speech band is entirely obliterated by noise. The current use of the half-way point on this psychometric function provides a sensitive measure of noise susceptibility that can be applied to any target speech band regardless of the difference between band-present and band-absent scores, and regardless of the absolute intelligibility values. In the current study, the SNR required to achieve this half-way point was determined for each band and compared to evaluate each band's susceptibility to noise.

Apparent from Fig. 5 are large differences in the noise susceptibilities of the 20 bands across the spectrum. In fact, the

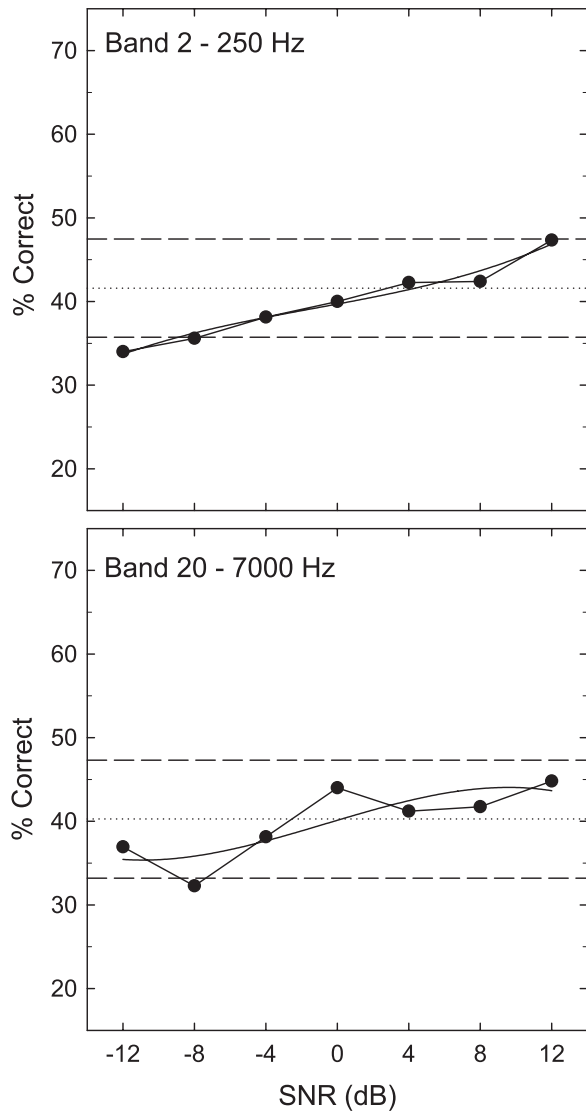


FIG. 4. Same as for Figs. 2 and 3, but data are from a new group of subjects for bands 2 and 20. The dashed band-present and band-absent scores were obtained from these same subjects.

difference in equivalent SNR between bands was found to be as large as 12.7 dB (250-Hz band versus 1170-Hz band). Thus, arguably the most important finding here is that noise susceptibility is not equal across the spectrum. Further, large differences occurred within a single region of the spectrum. Whereas the lower frequency bands (centered 450–1170 Hz) displayed relatively consistent noise susceptibility, the lowest band tested (250 Hz) differed considerably from this group. The higher frequency bands (1370 Hz and above) displayed large variability across bands, with the highest-frequency band (8500 Hz) being one of the most extreme values.

Accordingly, these data suggest that no simple pattern of noise susceptibility exists across the spectrum. This is true despite the fact that these data were obtained using a speech corpus of 10 different talkers (half of each gender). Therefore, these results do not simply reflect any particular acoustic idiosyncrasy of an individual talker, but rather the differences observed here are likely to be more global or generalizable across talkers.

TABLE II. Noise susceptibility (in equivalent dB SNR) and the difference from mean noise susceptibility (−1.95 dB SNR) for each speech band. Values for bands 2 and 20 are from the control conditions.

Band	center frequency (Hz)	SNR(dB)	SNR difference from mean (dB)
2	250	4.3	6.3
3	350	0.4	2.4
4	450	−4.0	−2.1
5	570	−1.6	0.4
6	700	−3.8	−1.9
7	840	−7.2	−5.3
8	1000	−3.2	−1.3
9	1170	−8.4	−6.5
10	1370	−0.6	1.4
11	1600	−4.0	−2.1
12	1850	1.7	3.7
13	2150	−4.4	−2.5
14	2500	−3.4	−1.5
15	2900	2.8	4.8
16	3400	1.8	3.8
17	4000	−1.0	1.0
18	4800	−3.0	−1.1
19	5800	2.4	4.4
20	7000	0.2	2.2
21	8500	−8.0	−6.1

Despite that the pattern is not simple, overall trends are observable. The smoothed data in Fig. 5 were prepared to examine these trends in noise susceptibility across the speech spectrum. The low frequencies contain a broad region of low susceptibility that reaches toward the center of the speech spectrum at 1500 Hz. Above that appears a region of average susceptibility (one approximately matching the dashed-line mean in Fig. 5), which is followed by a region of high susceptibility in the high frequencies. The exceptions are the very lowest and highest speech frequencies, which are some of the most and least susceptible to noise, respectively.²

As indicated in Sec. I, Studebaker and Sherbecoe (2002) created intensity-importance functions for different speech bands. Their interest was in characterizing speech band importance across a vast intensity range and comparing the resulting intensity functions to that contained in the SII. Speech was presented at 19–91 dB sound pressure level (SPL) in a 44 dB SPL speech-shaped noise, producing SNRs of −25 to 47 dB. Overall, it was found that speech contributions changed over a dynamic range somewhat greater than the 30 dB suggested by the SII. But it was also found that different speech bands contributed differently as a function of SNR. This finding might potentially reflect the differential noise-susceptibility concept examined currently.

The study of Studebaker and Sherbecoe (2002) possessed numerous and substantial differences in procedures and assumptions relative to the current study. But it may be possible to relate the five broad contiguous bands employed in the previous study to the broad regions of susceptibility observed currently in the smoothed data of Fig. 5. Specifically, Studebaker and Sherbecoe’s bands 2 (562–1122 Hz) and 5 (2818–8913 Hz) align roughly with current regions of low and high noise susceptibility, respectively. But correspondence is not readily apparent across the

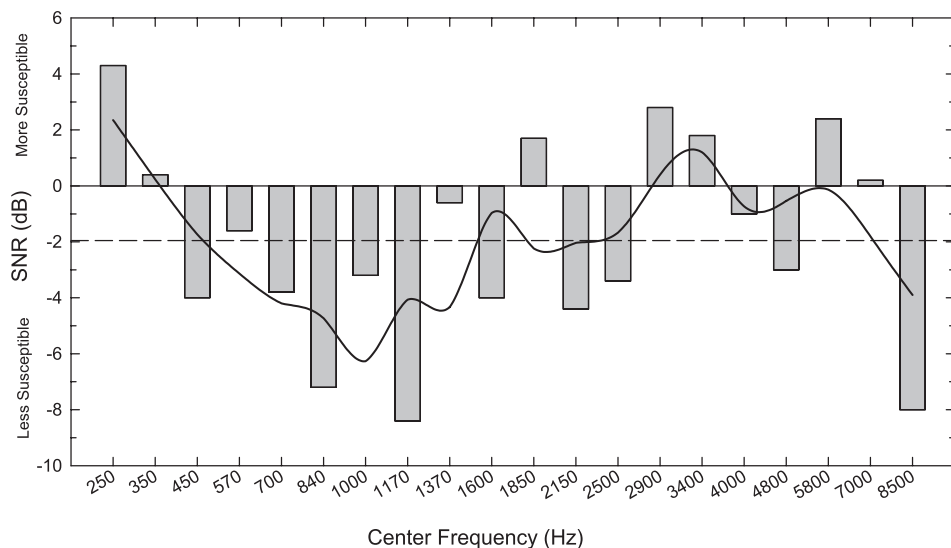


FIG. 5. Shaded columns show noise susceptibility as equivalent signal-to-noise ratios for 20 critical speech bands spanning 200–9500 Hz. Values for bands centered at 250 and 7000 Hz are from the control group shown in Fig. 4. The dashed line indicates average susceptibility across all 20 bands. The solid curve represents trends in noise susceptibility across frequency as equivalent signal-to-noise ratios. This curve was obtained by smoothing the individual values using a three-band rectangular sliding window and fitting with a spline curve.

results of the two studies, and band 5 can actually be interpreted as one of the least susceptible in the earlier work. Thus, direct comparison across the studies is difficult, and it is unclear to what extent the data reflect the same underlying mechanism.

We argue here that the susceptibility to noise that a speech band displays and its band importance are separate factors. There appears to be no systematic relationship between the two, even when compared using the same speech materials, recordings, and techniques. Figure 6 displays this lack of relationship through a scatterplot of band importance (values from Yoho *et al.*, 2018) versus noise susceptibility for each of the 20 bands tested. A Pearson’s correlation between importance and noise susceptibility was non-significant ($r = -0.24$, $p = 0.31$). By way of example, of the six bands showing the greatest susceptibility to noise in the current study, two were the least important bands examined, one had relatively low

importance, two had moderate importance, and one had the second-highest importance.

These findings illustrate a potential issue with evaluating speech band importance in background noise. As described in Sec. I, many current band-importance techniques confound importance with noise susceptibility. In fact, techniques that rely on altering the SNR to evaluate importance may in fact be measuring a band’s vulnerability to noise more than its importance. Other techniques that assess intelligibility at various cutoff frequencies in noise at different levels are likely measuring the combination of noise susceptibility and importance, because noisy speech bands contribute more to overall intelligibility if they are more important or if they are more robust to noise. One possible solution to this confound is the use of a technique that allows the examination of speech band importance in either quiet or in background noise. The technique employed here allows this flexibility. In previous work (Apoux and Healy, 2012; Healy *et al.*, 2013; Yoho *et al.*, 2018), speech band importance was examined for speech in quiet in order to isolate the effect of importance. This was possible by manipulating the number of “other” bands presented along with the target band in order to maintain performance in the steep portion of the psychometric function relating intelligibility to information present. However, if the desire is to evaluate band importance for speech in noise, or the combined effect of importance and susceptibility, the only necessary modification would be the addition of noise plus additional other band(s) as required to adjust overall performance depending on overall SNR employed.

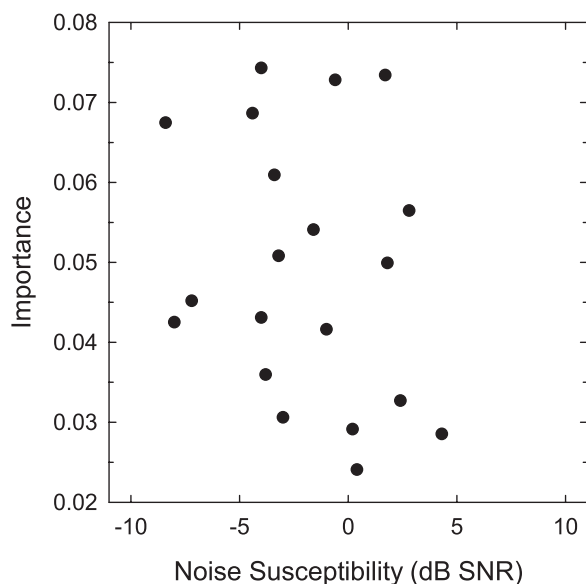


FIG. 6. Relationship between speech band importance (from Yoho *et al.*, 2018) and noise susceptibility in equivalent SNRs for 20 critical speech bands ($r = -0.24$, $p = 0.31$).

It may also be possible to incorporate the current concept of differing noise susceptibility across the speech spectrum into future versions of the SII. Currently, the assumption that noise impacts all speech bands similarly can be found in the constant relationship across bands between the SII audibility factor (A) and SNR in the band. The incorporation of noise susceptibility would be a simple matter of varying the relationship between (A) and SNR for each band, or by weighting (A) for each band.

Another possible application of the current technique involves the evaluation of the effect of sensorineural hearing

impairment on the relative importance of various regions of speech. The interactions on speech intelligibility between the effects of noise and broadened auditory tuning can be quite complex. Accordingly, it may be particularly advantageous to evaluate speech band importance and noise susceptibility separately for this population.

V. SUMMARY AND CONCLUSIONS

In the current study, the susceptibility to noise of various critical bands of speech was examined. The amount of noise within the band of interest was systematically varied to determine the SNR at which the intelligibility contribution of that band dropped by half. A multi-talker sentence corpus was employed so that noise susceptibility could be examined more generally, without the potential influence of one particular talker's voice.

- (i) In sharp contrast to common assumption and to what is currently implemented in the ANSI SII, the noise susceptibility of individual speech critical bands was found to vary greatly across the frequency spectrum.
- (ii) The range of noise levels (SNRs) required to affect various speech bands equivalently varied by nearly 13 dB, from as low as -8.4 dB SNR for the 1170-Hz band to as high as 4.3 dB SNR for the 250-Hz band.
- (iii) The observed pattern of noise susceptibility across frequency was not simple (see columns in Fig. 5)
- (iv) However, trends in noise susceptibility (see curve in Fig. 5) suggest that the bottom half of the speech spectrum is relatively robust to noise, whereas the upper half of the speech spectrum is average-to-more susceptible to the deleterious influence of noise. The very lowest and highest speech frequencies appear as exceptions to this pattern.
- (v) No relationship was observed between noise susceptibility and speech band importance.
- (vi) Incorporation of noise susceptibility into future versions of the SII would be computationally simple.
- (vii) The current results call into question the common practice of evaluating the relative importance of various bands of speech (the derivation of band-importance functions) in the presence of background noise. We argue here that susceptibility to noise and band importance are very different factors. These factors potentially play different roles in intelligibility, which may be underappreciated in previous work.

ACKNOWLEDGMENTS

This work was drawn from a dissertation submitted to the Ohio State University Graduate School by S.E.Y., under the direction of E.W.H. It was supported in part through grants from the National Institute on Deafness and other Communication Disorders (R01 DC008594 and R01 DC015521 to E.W.H.) and from the Ohio State University

Alumni Grants for Graduate Research and Scholarship. We are grateful for the data-collection and analysis assistance of Brittney Carter, Jordan Vasko, and Shuang Liu.

¹Due to an error in the processing script, three of the 40 subjects heard a small number of sentences twice. This corresponded to at most 1.3 sentences/condition. Because only a small minority of subjects were involved, the number of repeated sentences was very low, and the repeated sentences were randomly distributed across conditions, these data were retained.

²The reason for the exceptional noise susceptibility of the lowest and highest frequency bands is unclear, but the special significance of the most extreme frequencies has been recognized previously (e.g., Studebaker and Sherbecoe, 2002).

- ANSI (1969). ANSI S3.5. *American National Standard Methods for the Calculation of the Articulation Index* (American National Standards Inst., New York).
- ANSI (1997). ANSI S3.5 (R2007). *American National Standard Methods for the Calculation of the Speech Intelligibility Index* (American National Standards Inst., New York).
- ANSI (2004). ANSI S3.21 (R2009). *American National Standard Methods for Manual Pure-Tone Threshold Audiometry* (American National Standards Inst., New York).
- ANSI (2010). ANSI S3.6-2010. *American National Standard Specification for Audiometers* (American National Standards Inst., New York).
- Apoux, F., and Bacon, S. P. (2004). "Relative importance of temporal information in various frequency regions for consonant identification in quiet and in noise," *J. Acoust. Soc. Am.* **116**, 1671–1680.
- 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**, 1078–1087.
- Bacon, S. P., and Liu, L. (2000). "Effects of ipsilateral and contralateral precursors on overshoot," *J. Acoust. Soc. Am.* **108**, 1811–1818.
- 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**, 3769–3773.
- French, N. R., and Steinberg, J. C. (1947). "Factors governing the intelligibility of speech sounds," *J. Acoust. Soc. Am.* **19**, 90–119.
- Hawkins, J. E., Jr., and Stevens, S. S. (1950). "The masking of pure tones and of speech by white noise," *J. Acoust. Soc. Am.* **22**, 6–13.
- Healy, E. W. (1998). "A minimum spectral contrast rule for speech recognition: Intelligibility based upon contrasting pairs of narrow-band amplitude patterns," Ph.D. dissertation, The University of Wisconsin–Milwaukee; available from: <http://www.proquest.com/>; Publication Number: AAT 9908202, pp. 56–73.
- Healy, E. W., Yoho, S. E., and Apoux, F. (2013). "Band importance for sentences and words reexamined," *J. Acoust. Soc. Am.* **133**, 463–473.
- IEEE (1969). "IEEE recommended practice for speech quality measurements," *IEEE Trans. Audio Electroacoust.* **17**, 225–246.
- Miller, G. A. (1947). "The masking of speech," *Psychol. Bull.* **44**, 105–129.
- Miller, G. A., and Nicely, P. E. (1955). "An analysis of perceptual confusions among some English consonants," *J. Acoust. Soc. Am.* **27**, 338–352.
- Studebaker, G. A., Pavlovic, C. V., and Sherbecoe, R. L. (1987). "A frequency importance function for continuous discourse," *J. Acoust. Soc. Am.* **81**, 1130–1138.
- Studebaker, G. A., and Sherbecoe, R. L. (1991). "Frequency-importance and transfer functions for recorded CID W-22 word lists," *J. Speech Hear. Res.* **34**, 427–438.
- Studebaker, G. A., and Sherbecoe, R. L. (2002). "Intensity-importance functions for bandlimited monosyllabic words," *J. Acoust. Soc. Am.* **111**, 1422–1436.
- Webster, J. C., and Klumpp, R. G. (1963). "Articulation index and average curve-fitting methods of predicting speech interference," *J. Acoust. Soc. Am.* **35**, 1339–1344.
- Yoho, S. E., Healy, E. W., Youngdahl, C. L., Barrett, T. S., and Apoux, F. (2018). "Speech-material and talker effects in speech band importance," *J. Acoust. Soc. Am.* **143**(3), 1417–1426.