ASHA American Speech-Language-Hearing Association

JSLHR

Research Note

Updating the Spectral Correlation Index: Integrating Audibility and Band Importance Using Speech Intelligibility Index Weights

Gregory M. Ellis[a] and Pamela Souza[a,b]

[a] Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL [b] Knowles Hearing Center, Northwestern University, Evanston, IL

ARTICLE INFO

Article History:
Received August 18, 2021
Revision received December 16, 2021
Accepted April 3, 2022

Editor-in-Chief: Peggy B. Nelson
Editor: Yi Shen

https://doi.org/10.1044/2022_JSLHR-21-00448

ABSTRACT

The original Spectral Correlation Index (SCI_o) is a measure of amplitude envelope distortion that has been used in several studies to predict behavioral results. Because the original SCI_o did not account for the differential contribution of particular frequency bands to speech intelligibility (i.e., band importance) or for audibility, a new "individual" version (the SCI_i) is proposed and evaluated. Sentence intelligibility data are used to compare the predictive power and goodness-of-fit for statistical models using two versions of the SCI. The SCI_i provides significantly better fits to behavioral data than the SCI_o. This result demonstrates the importance of accounting for and including signal audibility in analyzing and modeling data collected from the population of individuals with hearing impairment. With this update, the SCI_i is a useful measure for predicting speech intelligibility based on amplitude envelope distortions.

The original Spectral Correlation Index (SCI_o) is a measure of amplitude envelope distortion that was proposed as a straightforward quantification of differences in amplitude modulation (AM) properties between a baseline and a comparison signal (Gallun & Souza, 2008). The SCI_o is obtained by calculating the modulation index within octave-wide bands centered on the audiometric frequencies (250, 500, 1000, 2000, 4000, and 8000 Hz). These are calculated for two signals: a baseline signal (typically an unprocessed speech signal in quiet) and a comparison signal (a signal that has undergone some amount of acoustic change: hearing aid compression, been subjected to background noise, frequency lowering, etc.). The modulation indices within octave bands for each signal are then vectorized and correlated with one another using a simple Pearson correlation. This results in a single measure that ranges from −1 to 1, but negative values tend to be rare. An SCI_o value of 1 represents a comparison signal whose AM properties are identical to the baseline. An SCI_o value of 0 represents a signal whose AM properties are uncorrelated with those of a baseline signal. Finally, an SCI_o value of −1 represents a signal whose AM properties vary inversely to the baseline.

A major advantage of the SCI_o is that it is capable of quantifying differences between two signals of different lengths. Other measures (like the Envelope Distortion Index and Hearing Aid Speech Perception Index and Hearing Aid Speech Quality Index) are limited in that they require the baseline and comparison signals to be of equal length. Because such measures require equal-length signals, it is impossible to quantify distortion due to processing strategies that change length of a signal (like time compression [TC]). The SCI_o does not have this limitation and, therefore, is worth iterating and improving upon.

The SCI_o has been shown to predict vowel–consonant–vowel (VCV) token perception; specifically, higher SCI_o values—indicating that modulation properties are preserved—are associated with fewer errors in VCV perception (Gallun & Souza, 2008; Souza et al., 2021; Souza & Gallun, 2010). However, the SCI_o as originally proposed is a somewhat generic measure. It does not account for the fact that certain carrier bands contribute more to speech understanding than others.

Correspondence to Gregory M. Ellis: gregory.ellis.phd@gmail.com.
Disclosure: *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

Such a band importance function is a key part of calculating both the Speech Transmission Index (Steeneken & Houtgast, 1980) and the Speech Intelligibility Index (SII; American National Standards Institute [ANSI], 1997), measures that provide accurate predictions of speech intelligibility. In addition, the SCI_o does not take into account band audibility relative to the listener's hearing thresholds. This research note proposes to combine the band importance function of the SII, audibility, and the distortion sensitivity of the SCI_o to improve the overall predictive power of the SCI_o and update it for predicting sentence materials rather than VCVs. An updated version of the SCI_o, the "individual SCI" (SCI_i), is proposed and compared against the original. This research note reanalyzes behavioral data that were published previously (Souza et al., 2021). The listeners, procedure, and stimuli are identical to that paper and reported here again for ease of reading and clarity. The goal of this paper is to quantify and discuss the differences in predictive power between the SCI_o and the SCI_i.

Method

Listeners

The listener data reported here were drawn from Souza et al. (2021). In that data set, 25 adults (14 women) with ages ranging from 63 to 89 years (M_{age} = 74 years) and bilateral sensorineural hearing loss participated. The better ear of each participant was chosen as the test ear. All participants passed the Montreal Cognitive Assessment (Nasreddine et al., 2005) with a score of at least 23 out of a possible 30 points (Shen et al., 2016). All listeners spoke American English as their primary language, completed an informed consent process, and were compensated for their time. All experimental methods were approved by the Northwestern University Institutional Review Board.

Stimuli

The speech signals were the same as those used in Souza et al. (2021). Test stimuli were Institute of Electrical and Electronics Engineers sentences (Rothauser, 1969). Two male and two female talkers were used. Talker selection and recording details are described in Panfili et al. (2017). Following from the experimental question for Souza et al. (2021) the sentences were time compressed, mixed with six-talker babble from the connected speech test (Cox et al., 1987) at 10 dB signal-to-noise ratio (SNR), and then passed through a hearing aid simulator.

TC was applied using Pitch Synchronous Overlap and Add (Moulines & Charpentier, 1990). This manipulation was performed in Praat (Boersma & Weenink, 2018) for three different TC rates. Sentences were compressed to 100% (no TC), 70%, and 50% of their original length. Pilot testing demonstrated these rates to produce behavioral results with no floor or ceiling effects.

Hearing aid processing was achieved using a hearing aid simulator (Kates, 2017) with a 16-channel filterbank. The center frequencies of the filterbank were 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 2, 2.5, 3, 4, 5, and 6 kHz. The first and last filters were low- and high-pass, respectively. All other filters were bandpass. Four wide dynamic range compression (WDRC) conditions were implemented: linear (1:1 compression ratio, 100 ms attack time, 1400 ms release time), slow (2:1 compression ratio, 100 ms attack time, 1400 ms release time), fast (2:1 compression ratio, 5 ms attack time, 100 ms release time), and hyper (4:1 compression ratio, 2 ms attack time, and 40 ms release time). After WDRC processing, gain was applied for each listener following the revised National Acoustics Laboratory gain rule (Byrne & Dillon, 1986). For more details on signal processing, please see Souza et al. (2021).

Procedure

The procedure for behavioral data is as reported in Souza et al. (2021). Testing was conducted in a double-walled sound-treated booth. Stimuli were converted from digital to analog using a TDT RX6 DAC, attenuated using a TDT PA5, and passed into a TDT HB7 headphone buffer. Signals were presented to listeners via an ER-3 insert headphone.

Sentences were blocked by TC condition. Uncompressed signals were presented first, followed by sentences that were 70% of the original length, then ending with sentences that were 50% of the original length. Each list contained 100 unique sentences. Sentence, talker, and WDRC condition were randomized within blocks. After listening to each sentence, the listener verbally repeated the sentence. The number of correct keywords out of a possible 5 was recorded by an experimenter located outside of the sound-treated booth. No feedback was provided. Listeners received a break after each block.

Calculating the SCI

The SCI_o is described in Gallun and Souza (2008). The steps of the SCI_i are detailed below and in Figure 1. All processing was done in MATLAB (The Math Works, 2021). Two signals are required to calculate any version of the SCI: a baseline signal (typically not processed by a hearing aid, or minimally processed) and a comparison signal. The sampling frequencies of the signals presented here were 44.1 kHz.

All processing steps are done for both the baseline and comparison signals. First, a signal is passed through a

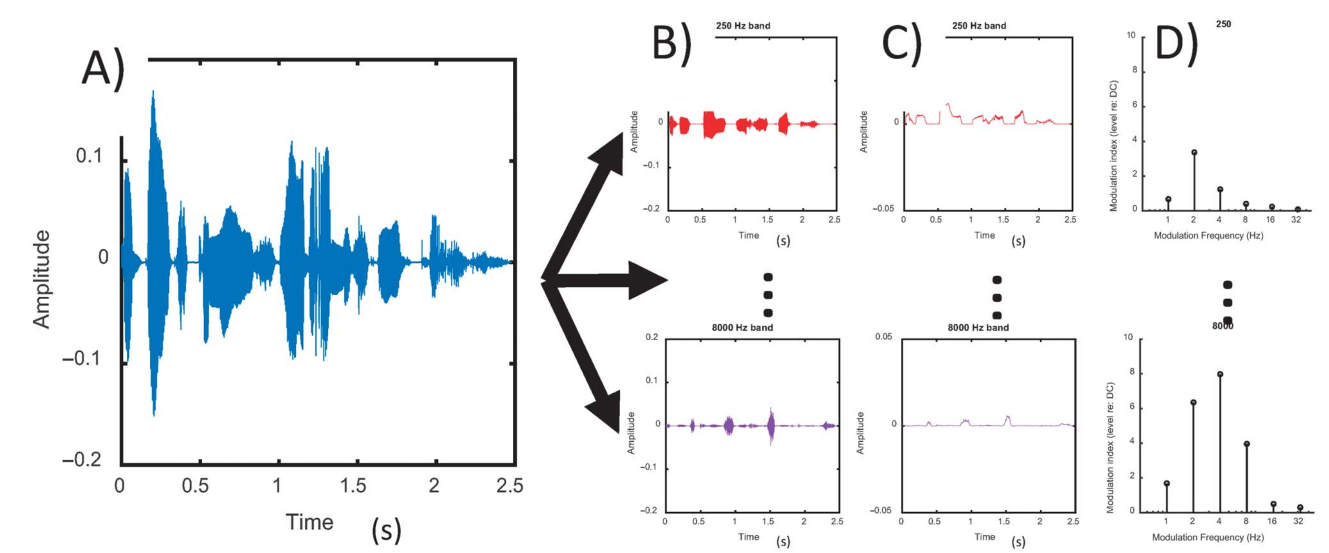


Figure 1. The first four steps in the Spectral Correlation Index (SCI) processing chain. A) Sample waveform taken from the Institute of Electrical and Electronics Engineers corpus. B) The signal is passed through a filterbank. Each filter is one octave wide with center frequencies equal to audiometric frequencies (octave frequencies between 250 and 8000 Hz). C) The output of the filters in B are low-pass filtered at 50 Hz. D) Fast Fourier Transform is performed on the resultant envelopes from C. The energy is binned in octave wide amplitude modulation (AM) frequencies between 1 and 32 Hz. The power in each bin is divided by the DC component (0 Hz). This provides the modulation index for a particular AM frequency within a particular carrier frequency.

bank of bandpass filters with center frequencies at 250, 500, 1000, 2000, 4000, and 8000 Hz. Next, the envelopes within each band are extracted. Envelope extraction is accomplished by half wave rectifying the outputs of the filters (setting all negative values to 0). The rectified signals are then processed with a finite impulse response low-pass filter (passband: 50 Hz; stopband: 250 Hz; stopband attenuation 60 dB). Then, the AM properties of each envelope are analyzed using Fourier analysis. In the SCI_o (Gallun & Souza, 2008), the signals were analyzed with a frequency resolution of 0.2 Hz. To keep the same frequency resolution, envelopes are resampled to a rate that would ensure an AM resolution of 0.2 Hz. The frequency resolution of a fast Fourier Transform (FFT) is defined as:

$$f_R = \frac{f_S}{N} \tag{1}$$

where f_R is the frequency resolution, f_S is the sampling frequency, and N is the number of points in the FFT. We can, therefore, determine the necessary f_S to attain a frequency resolution of 0.2 Hz as long as the length of the signal (N) is known. Given this, f_S is determined by taking the product of the length of the signal (N) and 0.2 Hz (e.g., a signal with 1000 points would be resampled to an f_S of 200 Hz). In the case that the resulting f_S is not possible due to limitations in MATLAB (i.e., the product of the new and old sample rates is greater than 2^{31}), the maximum fs that does not exceed the limitation in MATLAB is used instead.

In order to analyze the AM properties of each band, the output of the FFT is then subjected to a binning procedure. In the simplest terms, the output of the FFT is processed by a bank of perfectly rectangular filters with octave-wide passbands centered on 1, 2, 4, 8, 16, and 32 Hz. In practice, this is done by summing the energy that fell within the passbands of these rectangular filters. The output of these rectangular filters is divided by the energy of the DC component (0 Hz). Dividing by the energy present in the DC component provides a normalized "modulation index" value that indicates the relative amount of modulation within each AM filter and within each frequency band.

The result of this process is a 6 × 6 modulation index matrix (MIM) that contains the relative amount of modulation in each carrier frequency band (rows) and at each of the AM frequencies (columns). For the original SCI, the matrix was as described above. For the proposed new version, the MIM is subjected to a weighting procedure. The weighting procedure is based on an individual's hearing loss and the gain they receive at each frequency and is, therefore, named the "individual" SCI (SCI_i). The SCI_i provides a measure of distortion that is dependent on the listener's individual hearing abilities. Two components of the SII play a prominent role in calculating the SCI_i: the band audibility function (A) and the band importance function (I).

A mathematically describes the audibility of speech sounds within the bands centered on 250, 500, 1000, 2000, 4000, and 8000 Hz. A is calculated by solving several equations in sequence. These steps are detailed in the SII ANSI standard (ANSI, 1997) but are reproduced here for convenience and clarity. The first step in calculating A is the summation of an individual's thresholds with internal noise spectrum level to produce the equivalent internal noise spectrum level. This is done by summating the vector of internal noise spectrum (−3.9, −9.7, −12.5, −17.7, −25.9, −7.1) with an individual's effective hearing thresholds. The effective hearing thresholds are acquired by taking the difference between a listener's thresholds in dB HL and the gain applied by the hearing aid simulator, so these values differ from individual to individual. Next, the disturbance spectrum level (D) is calculated as the maximum between the SNR within a given band and the equivalent internal noise spectrum level. Then, the level distortion factor (L) is calculated, but for the stimuli examined here (in the terms of the SII standard, a "raised" vocal effort with 68.34 dB SPL) was always a vector containing only 1's. The next calculation is shown below:

$$K = (E - D + 15)/30 \tag{2}$$

where E is the equivalent speech spectrum level (here, 38.98, 40.15, 33.86, 25.32, 16.78, and 5.07 dB for each octave band) and D is the disturbance spectrum level calculated above. K is restricted to fall between 0 and 1, so values below 0 or above 1 are replaced with 0 or 1, respectively. K is then multiplied with L to produce A.

I mathematically describes the importance of those bands for understanding speech. The six-frequency I from the SII is defined as 0.06, 0.17, 0.23, 0.26, 0.21, and 0.05 for octave bands centered 250, 500, 1000, 2000, 4000, and 8000 Hz, respectively for average speech. These weights can (and should) be changed for different speech materials depending on the goal of the researcher (see Discussion for more consideration of this point).

Together, these functions capture two important aspects of speech understanding: audibility and importance. In order to integrate A and I into the SCI_i, the SCI_i is calculated by following many of the steps in calculating the SII. All of the steps of the SII are performed as laid out in the ANSI standard (ANSI, 1997) except for the summation in equation 14 from the standard. Rather, the product of A and I is used to weight the rows of the MIM. This weighting procedure is applied only to the comparison stimulus MIM. The MIMs are then vectorized and correlated. The following equation expresses these operations mathematically:

$$SCI_i = corr(vec(MIM_c \odot i), vec(MIM_b)) \tag{3}$$

where i is the product of the band audibility function (A) and band importance functions (I) from the SII standard (ANSI, 1997). Although the particular choice of the SII weights is somewhat arbitrary here (indeed, the weights are very similar to those of the Speech Transmission Index), the point of this step is to improve the predictability of the SCI by using information about hearing sensitivity for individual listeners. See Figure 2 for an example MIM, weighting function, and resulting weighted MIM.

The correlation can be defined more formally as follows:

$$SCI_i = \frac{\sum(c' - \bar{c'})(b - \bar{b})}{\sqrt{\sum(c' - \bar{c'})^2 \sum(b - \bar{b})^2}} \tag{4}$$

Where c' is the weighted comparison matrix, b is the unweighted comparison matrix, and the overbars represent the mean of the matrix. Equation 2 is a simple Pearson correlation.

Analysis and Results

Acoustic analysis using the SCI_o and the SCI_i was performed on the stimuli described above. The baseline stimulus was linearly amplified and uncompressed in time. Figure 3 shows the resulting distributions of SCI_o and SCI_i values (collapsed across all TC and WDRC conditions; for more details, see Souza et al., 2021) using violin plots. Violin plots visually summarize data by plotting

Figure 2. Example modulation index matrix, Speech Intelligibility Index (SII) weighting vector, and resultant weighted modulation index matrix.

Modulation Index Matrix (MIM)

Modulation Frequency (Hz); rows: Carrier Frequency (Hz)

	1	2	4	8	16	32
250	0.67	3.37	1.24	0.41	0.24	0.09
500	1.11	3.03	2.58	1.12	0.92	0.11
1000	5.81	4.01	6.91	1.98	1.18	0.26
2000	3.96	4.15	6.23	2.13	0.75	0.19
4000	0.66	1.11	2.46	3.82	1.2	0.59
8000	1.69	6.35	7.98	3.96	0.5	0.3

$\odot$

SII weights
0.0617
0.1671
0.2373
0.2648
0.2142
0.0549

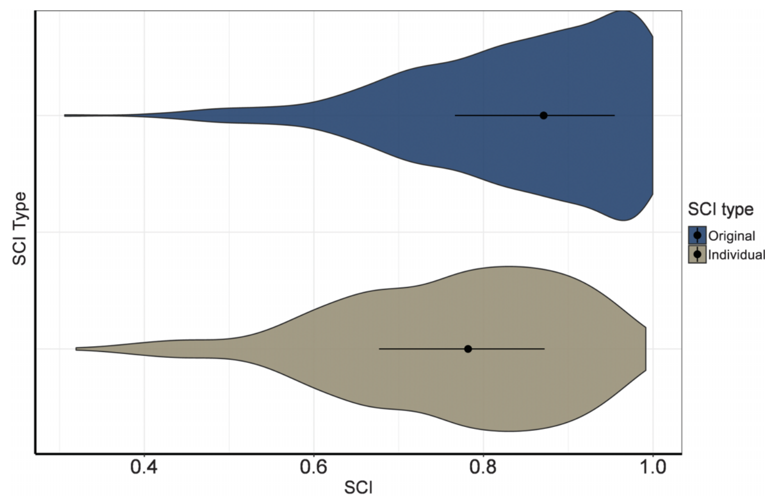
=

Weighted MIM

Modulation Frequency (Hz)

	1	2	4	8	16	32
250	0.04	0.21	0.08	0.03	0.01	0.01
500	0.19	0.51	0.43	0.19	0.15	0.02
1000	1.38	0.95	1.64	0.47	0.28	0.06
2000	1.05	1.10	1.65	0.56	0.20	0.05
4000	0.14	0.24	0.53	0.82	0.26	0.13
8000	0.09	0.35	0.44	0.22	0.03	0.02

Figure 3. Violin plots showing the distribution of Spectral Correlation Index (SCI) values for each of the SCI types. The original Spectral Correlation Index is plotted in blue (top) and the new "individual" version in gray (bottom). Black points represent medians of distributions. Black lines span the 25th and 75th percentile. The height of the "violin" represents the number of cases for a given SCI value and can be interpreted like a density plot or smoothed histogram. Both distributions tend to be negatively skewed; however, the individual SCI is less skewed than the original SCI.



frequency of occurrence (density) along the *y*-axis and the parameter of interest along the *x*-axis. Here, the distributions use the default kernel for calculating density in the ggplot package (Wickham, 2016) in R (R Development Core Team, 2021). Note that the distributions for the SCI_o and the SCI_i are qualitatively different from one another. The SCI_o is negatively skewed (skewness = −0.95), whereas the SCI_i is less negatively skewed (skewness = −0.61). The distribution of SCI_i values is more desirable for statistical modeling because it is closer to a normal distribution than the SCI_o.

All analyses were done in R (R Development Core Team, 2021) using the tidyverse package (Wickham et al., 2019) for data handling and visualization. A binomial generalized linear model (GLM) was used to assess whether either the individual SCI_i showed any marked improvement over the SCI_o. The dependent variable of the binomial GLM was percent correct on the speech intelligibility task. Independent variables were SCI type (a factor with two levels) and SCI value (a continuous variable, centered before analysis). A significant interaction between SCI type and SCI value would indicate a difference between the predictors, so the interaction term was included in the model.

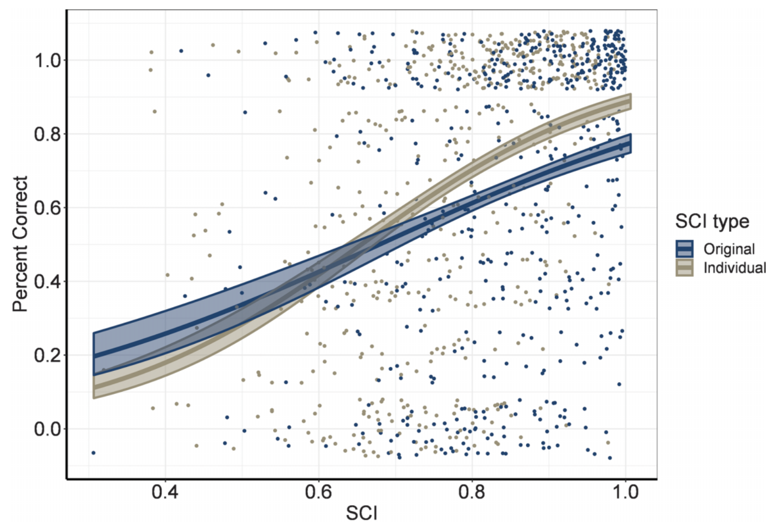
Overall, the model fit significantly better than chance, $\chi^2(3) = 468.2$, $p < .001$. There was a significant interaction between SCI type and SCI value, $\chi^2(1) = 20.09$, $p < .001$, as well as significant main effects of SCI type, $\chi^2(1) = 36.80$, $p < .001$, and SCI value, $\chi^2(1) = 411.32$, $p < 0.001$. The slopes between the individual SCI and the original SCI were significantly different ($\Delta\beta = 2.165$, $z = 4.467$, $p < .001$). See Figure 4 for a visualization of the fits.

While it is clear that the individual SCI leads to a different prediction, this does not necessarily tell us anything about the goodness-of-fit. To assess goodness-of-fit for both types of SCI, two separate binomial GLMs were fit to the intelligibility data for each SCI. Then, the Akaike information criterion (AIC) and deviation values of each model were compared. The AICs of the SCI_o and individual SCI_i models are 2449.03 and 2255.41, respectively. The AIC values on their own indicate that the individual SCI model is the best fit, though it can be difficult to interpret the magnitude of differences in AIC. To aid interpretation, the AIC values were transformed to AIC weights (Wagenmakers & Farrell, 2004). Briefly, AIC weights allow a direct comparison of how likely one model is relative to the next given the data examined (i.e., model A is twice as likely as model B, etc.). The AIC weights indicate that the model using the SCI_i is 43 *orders of magnitude* (7.33×10^{43}) times more likely than the $SCI_{o.}$ This suggests that the SCI_i is a significantly better model than the SCI_o.

Discussion

This research note presents an update to a method used to calculate the SCI_o Gallun & Souza, 2008) by using the band importance weights and the individualized band audibility function from the SII (ANSI, 1997). This updated distortion measure is referred to as the SCI_i. The SCI_i requires audiometric thresholds and applied gain information, but produces significantly stronger

Figure 4. Lines of best fit and 95% confidence intervals based on the different Spectral Correlation Index (SCI) measurements. Shaded regions represent the 95% confidence intervals. Points represent sentence scores. The original Spectral Correlation Index values are plotted in blue, whereas the Spectral Correlation Index, "individual" version (SCI_i) values are plotted in gray. Points are jittered along the *y*-axis to aid in visualization. It is clear that the SCI_i is more sensitive to changes in performance as a function of SCI value.



predictions (tens of orders of magnitude better) than the SCI_o by accounting for individual differences.

There are numerous benefits of the SCI over other distortion metrics regardless of the specific version of the metric. First, the SCI is capable of quantifying distortion between two signals that are not of equal length. This is an important advantage offered by the SCI. Although metrics like the Envelope Distortion Index (EDI) and Hearing Aid Speech Perception Index/Hearing Aid Speech Quality Index provide good predictions of behavioral data, when the signals of interest are not of equal length (i.e., TC), they cannot be used. The SCI solves this problem by examining signals in the AM frequency and carrier frequency domain, circumventing the need to have equal-length signals. This also provides a more nuanced approach to characterizing the AM properties of a signal than the EDI. The flexibility and nuance of the SCI are valuable properties of this metric.

This research note presents a reanalysis of previously published data (Souza et al., 2021) in which TC and WDRC parameters were systematically varied. In the present analysis, these data were handled fairly coarsely by collapsing across experimental condition (TC and WDRC parameters). Regardless of this coarse handling of the data, the SCI_i was still a strong predictor of performance. The strong predictive power of the SCI_i in even a coarse handling of data demonstrates that it is able to capture acoustic modifications produced by different experimental manipulations. For a by-condition analysis, see Souza et al. (2021).

There are several explanations for the improved predictive value of the SCI_i over the original SCI. In the SCI_o, all carrier bands were assumed to contribute equally to perception, which we know is not the case based on both the SII (ANSI, 1997) and the Speech Transmission Index (STI) (Steeneken & Houtgast, 1980). Weighting by band importance adjusts the metric to be consistent with principles of speech intelligibility. Also, in the SCI_o, contributions of modulation were included even if the carrier band of the modulations was below the listener's threshold. By not accounting for audibility, the SCI_o assumes all modulation information is fully available to the listener. This is not the case if a given carrier band is inaudible. The SCI_i addresses this by accounting for audibility and de-weighting modulations in carrier bands that are less audible to the listener according to the band importance function of the SII. While the weights used here do not represent the only possible modification, they are a principled approach to incorporating signal audibility and modulation in a straightforward metric.

In the future, it may be feasible to further customize the metric by considering the listener's modulation sensitivity, as well as sensitivity to sound level. Refinements could be made to the measure by narrowing the bandwidth of the carrier or modulation frequency filterbanks. Finally, while several options for suitable weights (e.g., SII, STI) are available, they are derived from average speech levels over time. Weighting functions that account for speech audibility in a more nuanced way than using the long-term average spectrum (e.g., Rhebergen & Versfeld, 2005; Shen et al., 2020) could also be informative.

Conclusions

The individual SCI (SCI_i) is an improvement over the original SCI (SCI_o) for calculating distortions caused by hearing aid processing like WDRC. The SCI_i utilizes individual gain and hearing loss information to achieve this improvement.

Acknowledgments

This work was funded by Grant NIDCD R01 DC006014, awarded to Pamela Souza. The authors thank Frederick (Erick) Gallun and Richard Wright for helpful comments on an earlier version of this research note.

References

American National Standards Institute. (1997). *Methods for calculation of the speech intelligibility index*. (ANSI S3. 5–1997 (R2017))

Boersma, P., & Weenink, D. (2018). *Praat: Doing phonetics by computer (Version 6.0.42)*. http://www.praat.org

Byrne, D., & Dillon, H. (1986). The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear and Hearing, 7*(4), 257–265. https://doi.org/10.1097/00003446-198608000-00007

Cox, R. M., Alexander, G. C., & Gilmore, C. (1987). Development of the Connected Speech Test (CST). *Ear and Hearing, 8*(Suppl. 5), 119S–126S. https://doi.org/10.1097/00003446-198710001-00010

Gallun, F., & Souza, P. (2008). Exploring the role of the modulation spectrum in phoneme recognition. *Ear and Hearing, 29*(5), 800–813. https://doi.org/10.1097/AUD.0b013e31817e73ef

Kates, J. (2017). *HASim (Version 4)*.

Moulines, E., & Charpentier, F. (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication, 9*(5–6), 453–467. https://doi.org/10.1016/0167-6393(90)90021-Z

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Panfili, L. M., Haywood, J., McCloy, D. R., Souza, P. E., & Wright, R. A. (2017). *The UW/NU Corpus (Version 2.0)*. https://depts.washington.edu/phonlab/projects/uwnu.php

R Development Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org

Rhebergen, K. S., & Versfeld, N. J. (2005). A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. *The Journal of the Acoustical Society of America, 117*(4), 2181–2192. https://doi.org/10.1121/1.1861713

Rothauser, E. H. (1969). IEEE recommended practice for speech quality measurements. *IEEE Transactions on Audio and Electroacoustics, 17*(3), 225–246. https://doi.org/10.1109/TAU.1969.1162058

Shen, Y.-J., Wang, W.-A., Huang, F.-D., Chen, J., Liu, H.-Y., Xia, Y.-L., Han, M., & Zhang, L. (2016). The use of MMSE and MoCA in patients with acute ischemic stroke in clinical. *International Journal of Neuroscience, 126*(5), 442–447. https://doi.org/10.3109/00207454.2015.1031749

Shen, Y., Yun, D., & Liu, Y. (2020). Individualized estimation of the Speech Intelligibility Index for short sentences: Test-retest reliability. *The Journal of the Acoustical Society of America, 148*(3), 1647–1661. https://doi.org/10.1121/10.0001994

Souza, P., Ellis, G., Marks, K., Wright, R., & Gallun, F. (2021). Does the speech cue profile affect response to amplitude envelope distortion? *Journal of Speech, Language, and Hearing Research, 64*(6), 2053–2069. https://doi.org/10.1044/2021_JSLHR-20-00481

Souza, P., & Gallun, F. (2010). Amplification and consonant modulation spectra. *Ear and Hearing, 31*(2), 268–276. https://doi.org/10.1097/AUD.0b013e3181c9fb9c

Steeneken, H. J., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *The Journal of the Acoustical Society of America, 67*(1), 318–326. https://doi.org/10.1121/1.384464

The Math Works. (2021). *MATLAB (Version 2021a)*. http://www.mathworks.com

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11*(1), 192–196. https://doi.org/10.3758/BF03206482

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Muller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686