

Supplemental Material S1. Statistical appendix.

Let TMR_{ij} be the “target-to-masker ratio,” defined as the presentation dB required by subject i under stimulus condition j to achieve 50% correct performance on the speech-on-speech masking task. TMR_{ij} is calculated according to the methods described in Gallun et al. (2015) by $TMR_{ij} = 10 - 20 \cdot p_{ij}$, where p_{ij} is the probability of a correct response by the i^{th} subject given the j^{th} stimulus condition. We model p_{ij} via multi-level logistic regression, with stimulus condition, diabetes group, low-frequency (.5, 1, 2 kHz) pure-tone average (PTA), low-frequency hearing asymmetry, high-frequency (3, 4, 6, 8 kHz) PTA, high-frequency asymmetry, age, age-by-stimulus condition interactions, diabetes group-by-stimulus condition interactions, and diabetes group-by-stimulus condition-by-age interactions. Let y_{ij} denote the number of correct responses out of 40 trials so that $y_{ij} \sim \text{Binomial}(p_{ij}, 40)$, and

Equation (1)

$$\begin{aligned} \log\left(\frac{p_{ij}}{1-p_{ij}}\right) = & \beta_0 + \beta_1 \cdot a_j + \beta_2 \cdot r_j + \beta_{a,r[j]} && \text{(main effects of stimulus condition)} \\ & + \delta_{D[i]}^0 + \delta_{D,a,r[ij]}^1 && \text{(main effects of diabetes group)} \\ & + (\theta_0^L + \theta_{a,r[j]}^L) \cdot LP_i && \text{(Low-freq PTA effects)} \\ & + (\gamma_0^L + \gamma_{a,r[j]}^L) \cdot LA_i && \text{(Low-freq asymmetry effects)} \\ & + (\theta_0^H + \theta_{a,r[j]}^H) \cdot HP_i && \text{(High-freq PTA effects)} \\ & + (\gamma_0^H + \gamma_{a,r[j]}^H) \cdot HA_i && \text{(High-freq asymmetry effects)} \\ & + (\alpha_0 + \alpha_{a,r[j]} + \alpha_{D,a,r[ij]}) \cdot Age_i && \text{(age and diabetes-specific age effects)} \\ & + \xi_i && \text{(residual subject-specific variation)} \end{aligned}$$

In this model a_j is the spatial separation of the masker (0° , 8° , or 30°) and r_j is a binary indicator for anechoic ($r_j = 0$) and reverberant ($r_j = 1$) settings. The predictors LP_i , LA_i , HP_i , and HA_i correspond to low-frequency PTA, low-frequency left-right ear asymmetry, high-frequency PTA, and high-frequency left-right ear asymmetry. The random effect ξ_i models subject-to-subject variation that is unmodeled by age or audiometric function.

The subscripting $a, r[j]$ denotes the spatial separation and reverberant setting for the ij^{th} response, and $D[i]$ denotes the diabetes group of subject i . The subscript $D, a, r[ij]$ denotes the diabetes group, spatial separation, and reverberant setting of the j^{th} stimulus presented i^{th} subject. Accordingly, the LP, HP, LA, HA effects are stimulus-condition specific, as indicated in equation (1). This notation conveniently specifies hierarchical main effects and interactions without resorting to matrix notation (Gelman & Hill, 2007).

TMR_{ij} is a linear function of p_{ij} . SRM is also a linear function of p_{ij} , and is written as $SRM_{i,r_j}^{a'} = TMR_{i,(a_j=0,r_j)} - TMR_{i,(a_j=a',r_j)}$, where a' is the spatial separation for which SRM is computed. For example, in the anechoic setting the SRM at 8° for subject i is $SRM_{i,(r_j=0)}^{8^\circ} = TMR_{i,(a_j=0^\circ,r_j=0)} - TMR_{i,(a_j=8^\circ,r_j=0)}$.

Priors. The intercept β_0 is given a Normal (.25,1) prior, reflecting the extensive literature with people varying in age and hearing loss published over the past five decades (e.g., Marrone et al., 2008b; Jakien & Gallun, 2018) suggesting that performance is generally quite good on the SRM task. The spatial separation main effect β_1 is given a Normal (.5,1) prior reflecting our

belief, based on this same literature, that performance improves as spatial separation increases. The parameters $\beta_2, \theta_0^L, \gamma_0^L, \theta_0^H, \gamma_0^H, \alpha_0$ are given weakly information Normal (0,1) priors, and all other effects are given hierarchical Normal (0, τ) priors, where τ is effect-specific and is given a half-normal (1) prior.

Computation. The model was fit in SAS Software version 9.4 with PROC Markov Chain Monte Carlo (MCMC). Three chains of the No-U-Turn sampler was run with random starting values for 2,000 iterations each. Convergence was indicated by Gelman-Rubin diagnostics all below 1.1. Posterior inferences are based on the 3 (chains) x 2,000 iterations = 6,000 posterior iterations.

Posterior Inference. We estimate the effects of diabetes, age, and hearing loss on TMR, and therefore SRM, by the method of “posterior predictive comparisons” (Gelman & Pardoe, 2007). Let $i(\mathbf{x})$ denote diabetes, age, and hearing inputs from the i^{th} subject, along with stimulus information. Similarly define $i(\mathbf{x}')$ where all inputs except one feature of interest, such as diabetes group or stimulus condition, are the same as in $i(\mathbf{x})$. The posterior predictive comparison is given by the difference in the model-based estimate of TMR (or SRM) with covariates set to $i(\mathbf{x})$ and set to $i(\mathbf{x}')$, i.e. $TMR_{i(\mathbf{x})} - TMR_{i(\mathbf{x}')}$ and by $SRM_{i(\mathbf{x})} - SRM_{i(\mathbf{x}')}$. For any particular contrast of interest, these differences are computed for each subject for each posterior MCMC sample. For example, to estimate the adjusted effect of uncontrolled DM2 on spatial release from masking at 30° in the anechoic condition, we compute equation (1) for each subject, inputting these stimulus conditions and each subject’s observed pure-tone threshold average and age, and setting diabetes group to “No DM.” This gives an estimate of TMR (or SRM) for that subject as if they had no diabetes. We again compute equation (1), but this time changing diabetes group to “Uncontrolled DM2,” corresponding to estimate of this subject’s performance as if they had uncontrolled DM2. The posterior predictive comparison is defined as the difference between these two estimates. This algorithm is repeated for each subject in the sample over each iteration of the MCMC sampler. Since diabetes group is the only feature that is varied, this algorithm gives the posterior distribution of the effect of uncontrolled DM, on TMR or SRM compared to No DM. Further details are found in Gelman and Pardoe (2007) and Gelman and Hill (2007).

The posterior distribution of the contrasts computed according to this algorithm describe our uncertainty about the effects of diabetes or age or hearing loss on speech understanding in complex listening environments, as measured using the SRM task. A point estimate of the contrast can be taken from the median of the posterior distribution. A 90% Bayesian confidence interval (called a “credible interval”) can be derived from the 5th and 95th percentiles of the posterior distribution. This interval reflects the range of effect sizes within which we are 90% certain the true value lies, given the data and the fitted model. Other intervals, such as the posterior inter-quartile range are similarly computed.