Supplemental Material S1.

As we mentioned in the manuscript, we analyzed different factors and parameters; however, we did not report part of them because we thought they did not add anything to the field. However, due to the request of the respected reviewers and for clarification of this study, we added some explanation as the supplementary documents as follows:

Global and Local Approaches

During data analysis, it became clear that the variability of the masker level, and hence the signal-to-noise ratio (SNR) of the target stimulus, was influencing the results. This analysis was based on both statistical analysis of frequency following response (FFR) SNR and the distribution of signal and noise for the entire length of an individual condition. To reduce the SNR range per condition, "local" stimulus SNRs were defined and specific ranges (low, medium, and high local SNRs) investigated separately. As defined in Figure 22, "Local," or short-term SNR, was calculated from target and masker powers taken during target presentation corresponding to each 250 ms time frame over which the target vowel occurred (Vondrasek & Pollak, 2005). Local SNRs are shown in relation to global SNRs in the top panel of Figure S1, below. The distributions were the same for all global SNRs, except for a shift of 5 or 10 dB. The three global SNRs were pooled together, and one distribution drawn, as depicted in the bottom half of Figure S1. The range of local SNRs was divided into three parts: a low, a medium, and a high SNR range, with the number of data points equal in each range. The ranges were as follows (in dB SNR): –12.9 to –2.4 (low), –2.4 to 3.2 (medium), 3.2 to 60.4 (high).

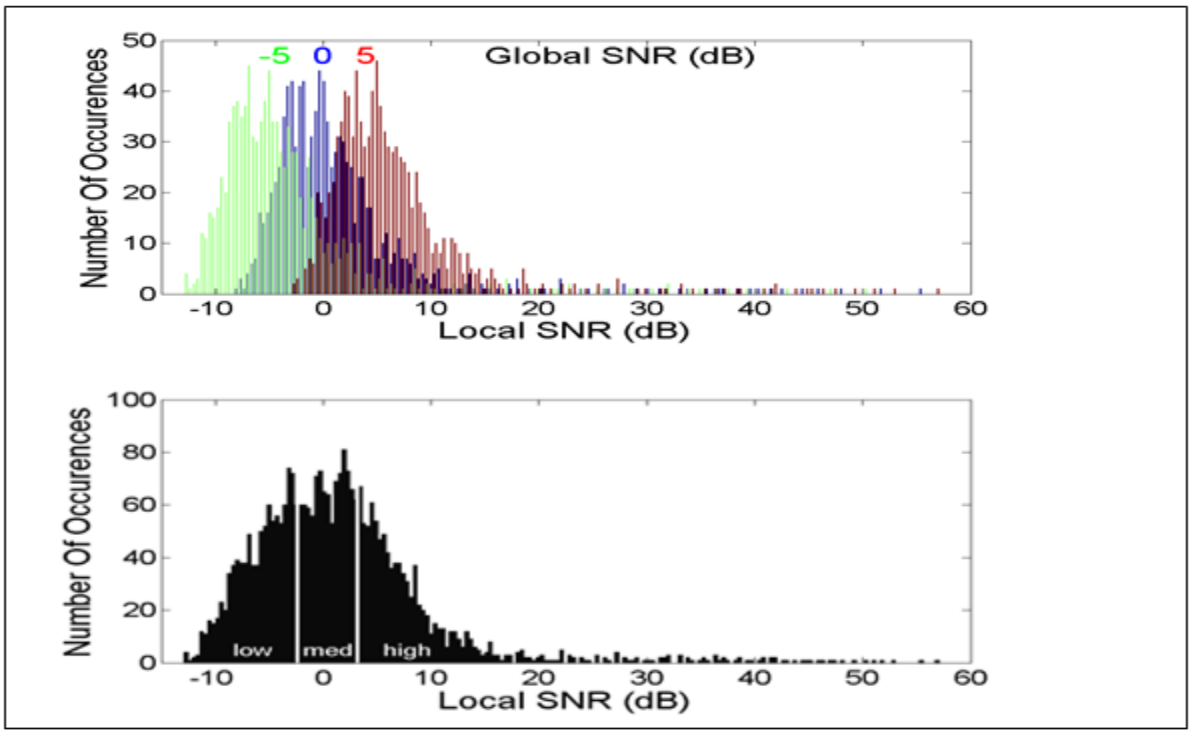

Figure S1. Low, medium, and high SNR ranges.

Higher Harmonics and Formants

To identify the strength of recorded signals relative to the recorded noise, a preliminary examination of the response SNRs for each FFR component was carried out, which prompted the focus of analysis on the fundamental frequency (F0) FFR component. The mean SNRs of the higher harmonics (2F0 to 5F0) and formants F1 and F2 ranged between –3 and 2 dB (Figure S2). The majority of these responses were not found to be significantly different from the noise floor, by comparing the F-ratio formed from the response power and the average power of the surrounding noise bins (Dobie & Wilson, 1996). Hence, it was decided to leave the higher harmonics and formants out of the analysis.

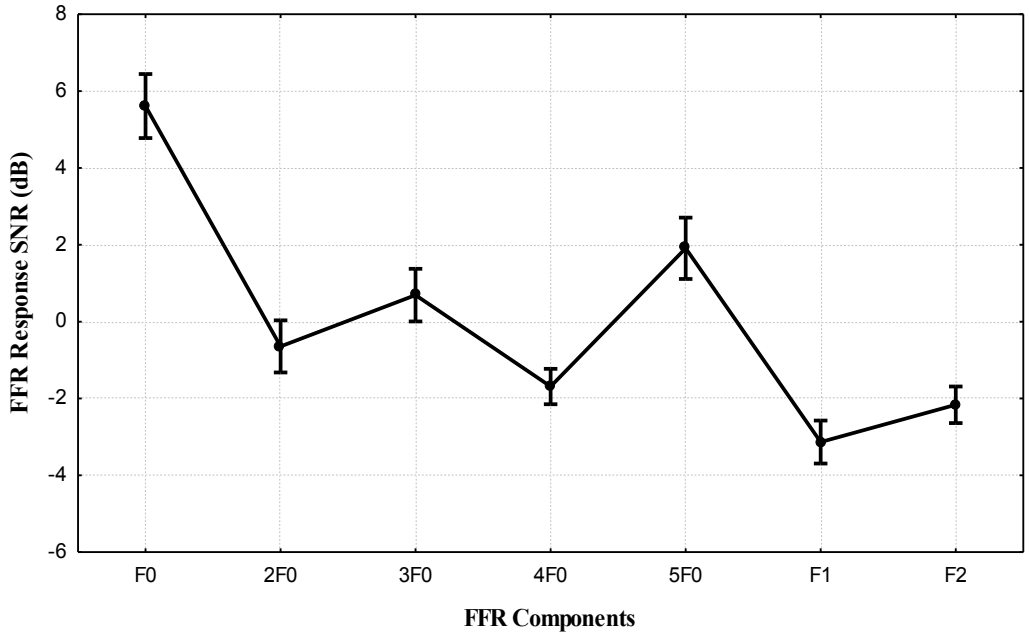

Figure S2. Means and standard errors of response SNR for each component of the FFR, averaged across recording channels.

Electrode Montage

Our analysis indicated that channel one (Fz-C7 montage) and channel three (Fz-A1&A2 montage), which were known as vertical montage, were similar in SNR (p = .052) and had significantly higher response SNRs than channel two (A2-A1 montage), which was known as the horizontal montage (p < .00001; Figure S3). As it is quite clear from Figure S3, the response SNR for Channel Two (the horizontal montage) was negative. It means that the collected data via this montage are mostly noise instead of signal. Therefore, all analyses were based on data collected through Channels One and Three.

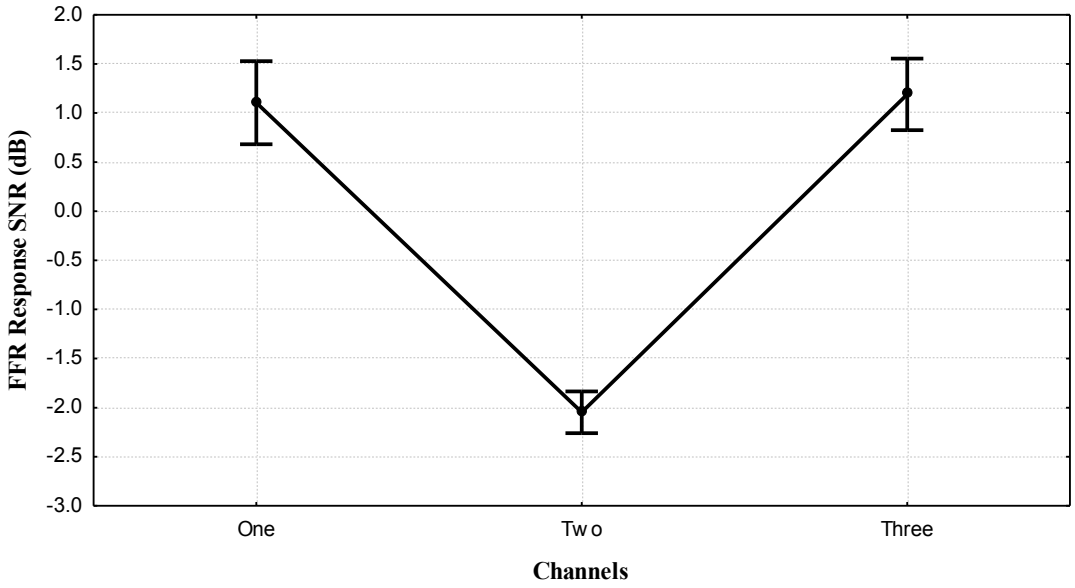

Figure S3. Means and standard errors of response SNR of FFR in recording channels, averaged across response components.

Channel Two (horizontal montage):

Figure S4 shows FFR F0 amplitudes for spatially co-located and separated conditions in horizontal montage. A significant effect of SRM was found on FFR amplitude ($F(1, 17) = 4.38$, $p = .05$). As it is crystal clear, we can see the same trend for SRM in Channel Two (horizontal montage) as in Channels One and Three (vertical montage). However, due to the negative response SNR in Channel Two, we cannot report these noise corrupted-data.

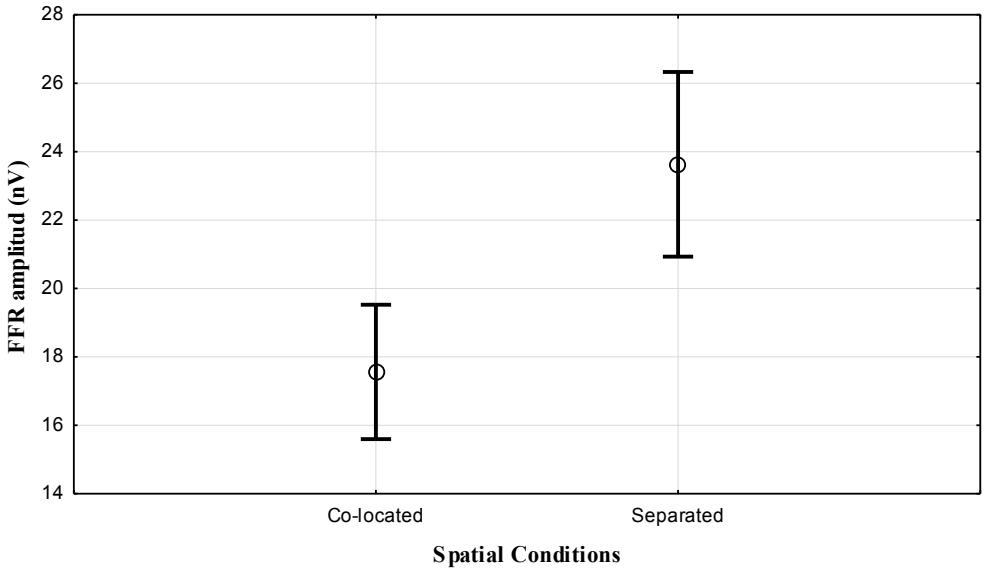

Figure S4. Means and standard errors of spatial co-located and separated of FFR F0 amplitude in channel one (horizontal montage), averaged across response components.

Figure S5 shows FFR F0 amplitudes for no-attention and attention conditions in channel two (horizontal montage). A significant effect of attention was found on FFR amplitude ($F(1, 17) = 12.62$, $p = .002$). Further confirmation for data we have collected in Channels One and Three. However, as we mentioned above, we cannot report these data due to negative response SNR in Channel Two (horizontal montage).

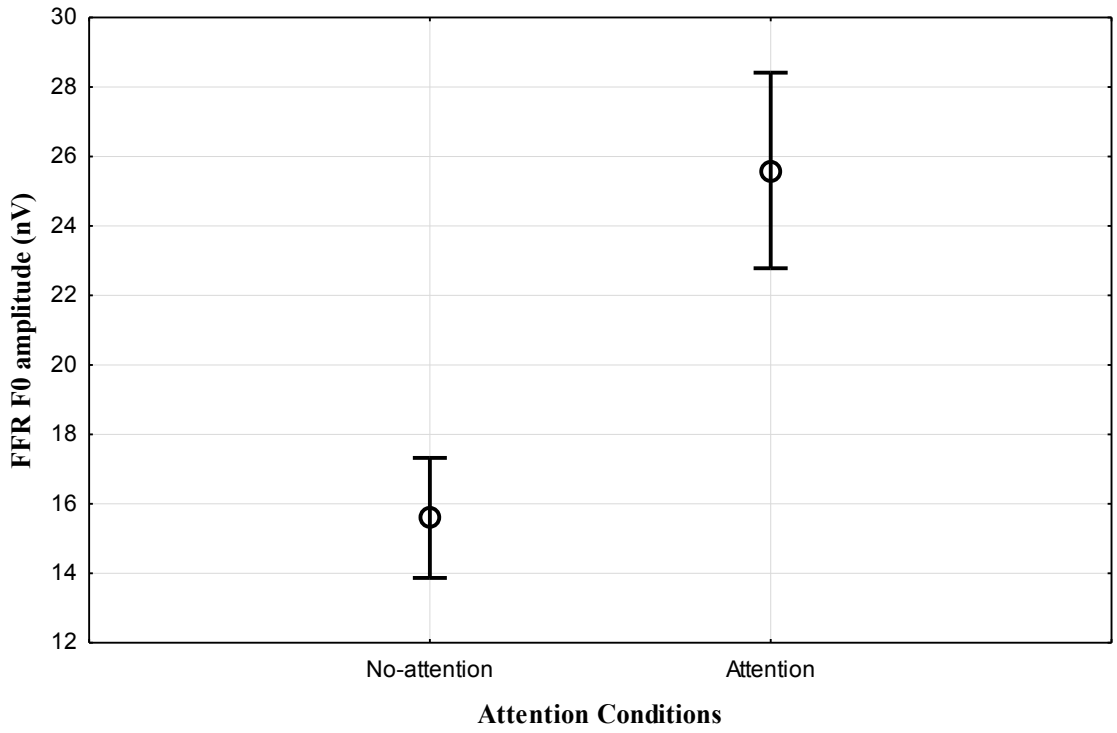

Figure S5. Means and standard errors of spatial co-located and separated of FFR F0 amplitude in channel one (horizontal montage), averaged across response components.

A brief review of results for global SNR analysis:

- **Evidence for spatial release from masking (SRM) in FFR F0 amplitude**
  - No significant main effect of SRM was found on FFR amplitude in local SNR analysis.
  - The lack of a main effect of SRM on F0 amplitude was confirmed when analysing the global SNR data as well ($F(1, 17) = 0.76$, $p = .40$).
- **Main effects of stimulus SNR**
  - With increasing stimulus SNR, the mean F0 amplitude in local SNR analysis became significantly larger ($F(2, 34) = 11.29$, $p = .00017$).
  - The findings were confirmed when analyzing the global SNR data as well. Mean F0 amplitudes increased significantly ($F(2, 34) = 37.69$, $p < .00001$) with increasing global stimulus SNRs (–5, 0, and 5 dB).
- **Interaction effects of SRM and SNR**
  - The interaction effect between SRM and stimulus SNR in local SNR analysis was significant ($F(2, 34) = 5.20$, $p = .011$).

  - A similar effect of SRM was found at –5 dB SNR in global SNRs (p=0.05), akin to the results for the low local SNR condition.
- **Main effects of attention**
  - A significant main effect of attention on the FFR F0 amplitude was found in local SNR analysis ($F(1, 17) = 8.96$, $p = .008$).
  - For global SNRs, a statistically significant effect of attention was found as well ($F(1, 17) = 7.42$, $p = .01$).
- **Interaction effects of attention and SRM**
  - No significant interactions effects were found between SRM and attention in local SNR analysis ($F(1, 17) = 0.03$, $p = .86$).
  - Analysis of global SNRs could not reveal any significant interaction effect between attention and SRM as well ($F(1, 17) = 0.03$, $p = .86$).

References

Dobie, R. A., & Wilson, M. J. (1996). A comparison of *t* test, *F* test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids. *The Journal of the Acoustical Society of America, 100*(4), 2236–2246. https://doi.org/10.1121/1.417933

Vondrasek, M., & Pollak, P. (2005). Methods for speech SNR estimation: Evaluation tool and analysis of VAD dependency. *RadioEngineering, 14*(1), 6–11. Retrieved from https://www.radioeng.cz/fulltexts/2005/05_01_06_11.pdf