Supplemental Materials

(Larrea-Mancera et al. 2022)

1. Effects of age and self-reported hearing ability on performance

Participants' scores on the Spanish version of the HHIA (instrument details provided below) had an average value of 3.37 (*SD* = 4.63; range 0-22). Overall, there were no statistically significant correlations between these scores and performance on the CRM or the the non-speech auditory assessments. Table S1 shows the correlations among the covariates of Age and HHIA score and each of the tests performed by participants in this study (average of sessions). We did not correct for multiple comparisons to better allow exploratory analysis potential effects to pursue in further studies. To formally analyze the potential influence of Age and HHIA scores we utilized both a composite score of all measures for each participant (Figure S1), and also, the assessment by assessment analysis (Figure S2). Figure S1 shows the composite across two sessions with a color code to illustrate the distribution of the variables of Age (left panel) and HHIA score (right panel). Non-significant correlations ($p \geq 0.27$) to the average of sessions confirm in a single result what can also be observed in the exploration of correlations to individual tests (see Figure S2). For this age group and people who have no self-reported hearing difficulties, there are no clear effects of Age nor HHIA on central auditory processing performance as measured with PART.

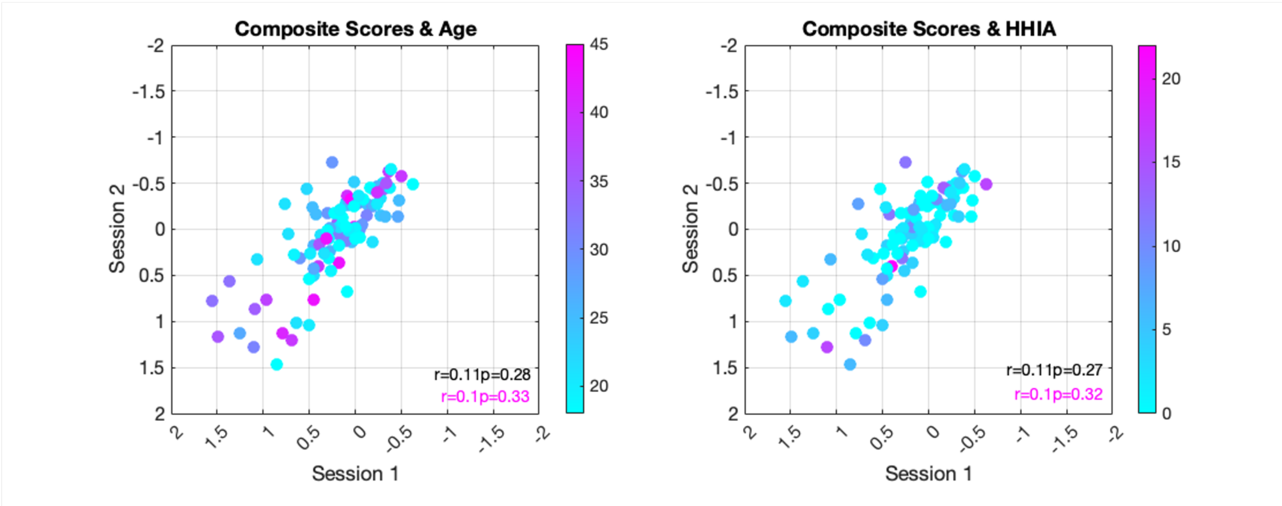

Figure S1. Scatter plots showing the strength of association of the covariates of Age (left) and HHIA score (right) indicated as a color-code (see color bar on the right of each figure). Stats in black indicate correlation to session 1 composite and stats in magenta indicate correlation to session 2 composite.

Table S1. Normative values (means and standard deviations) and correlations of each of the central auditory measures collected and the covariates of participant age and self-reported score on the HHIA.

Assessment	*Mean (sd)*	*Units*	*Age r (p)*	*HHIA r (p)*
Gap	3.01 (2.52)	ms	0.08 (.47)	0.101 (.32)
DichoticFM	0.67 (2.009)	Hz	0.27 (.01)*	0.015 (.88)
DioticFM	9.14 (1.57)	Hz	-0.06 (.54)	0.07 (.49)
TM	1.96 (0.78)	M dB	0.11 (.32)	0.05 (.62)
SM	1.68 (0.68)	M dB	-0.11 (.28)	0.03 (.74)
STM	1.06 (0.44)	M dB	0.02 (.83)	0.06 (.52)
No-Notch	-4.95 (1.66)	TMR dB	0.12 (.24)	0.17 (.09)
Notch	-17.52 (2.21)	TMR dB	0.1 (.33)	0.105 (.3)
Co-located	1.29 (1.16)	TMR dB	-0.04 (.7)	0.027 (.79)
Separated	-5.67 (2.68)	TMR dB	0.2 (.051)	0.04 (.68)
Spatial Release	6.96 (2.51)	dB	-0.23 (.023)*	-0.03 (.76)

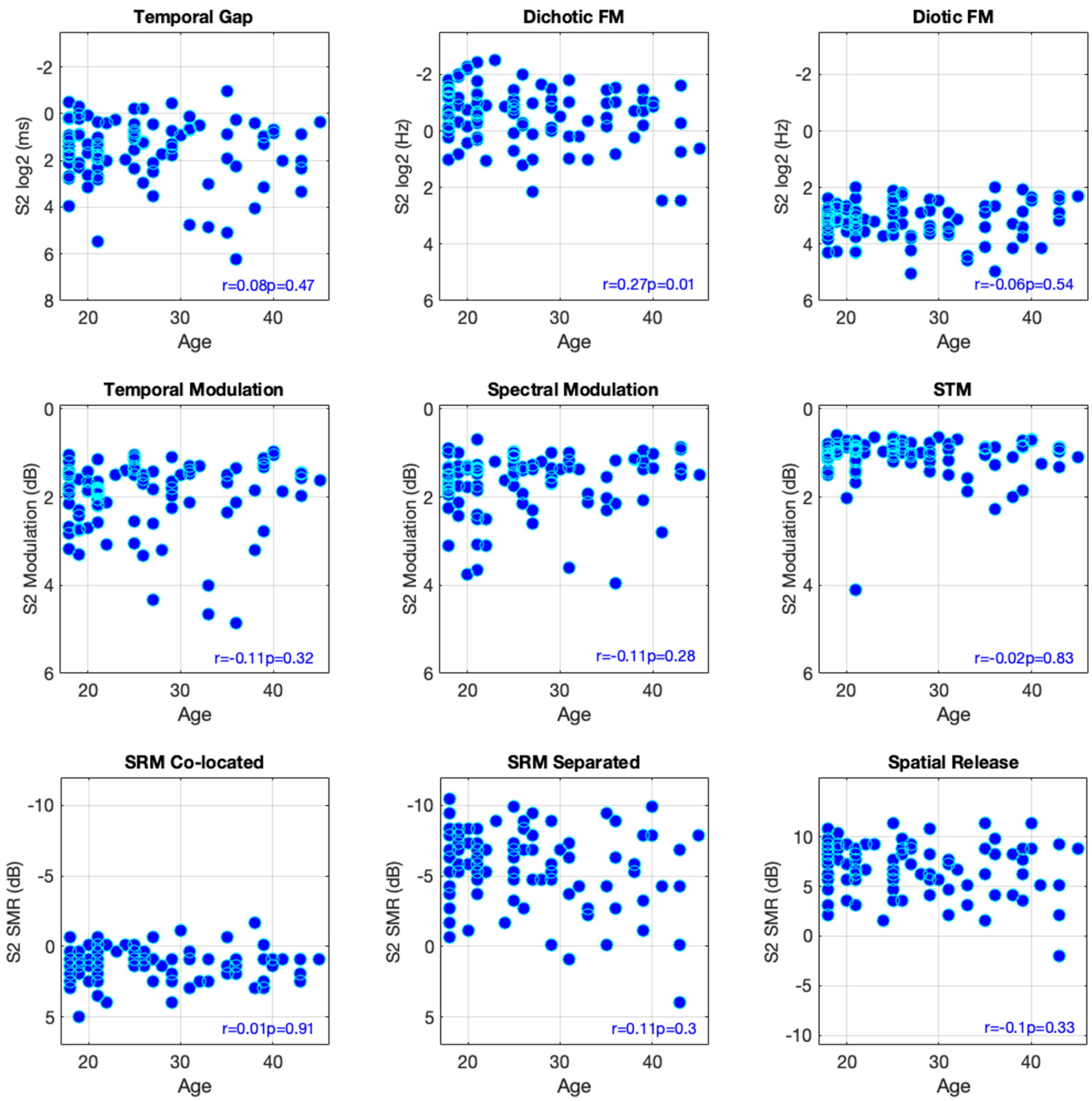

Figure S2. Scatter plots showing the strength of association of the covariate of Age (left) and the threshold obtained in each assessment. Correlation values are not significant, the corrected *p-value* for significance at *a* = .05 for this family of comparisons is .005.

2. Limits of agreement of non-speech measures

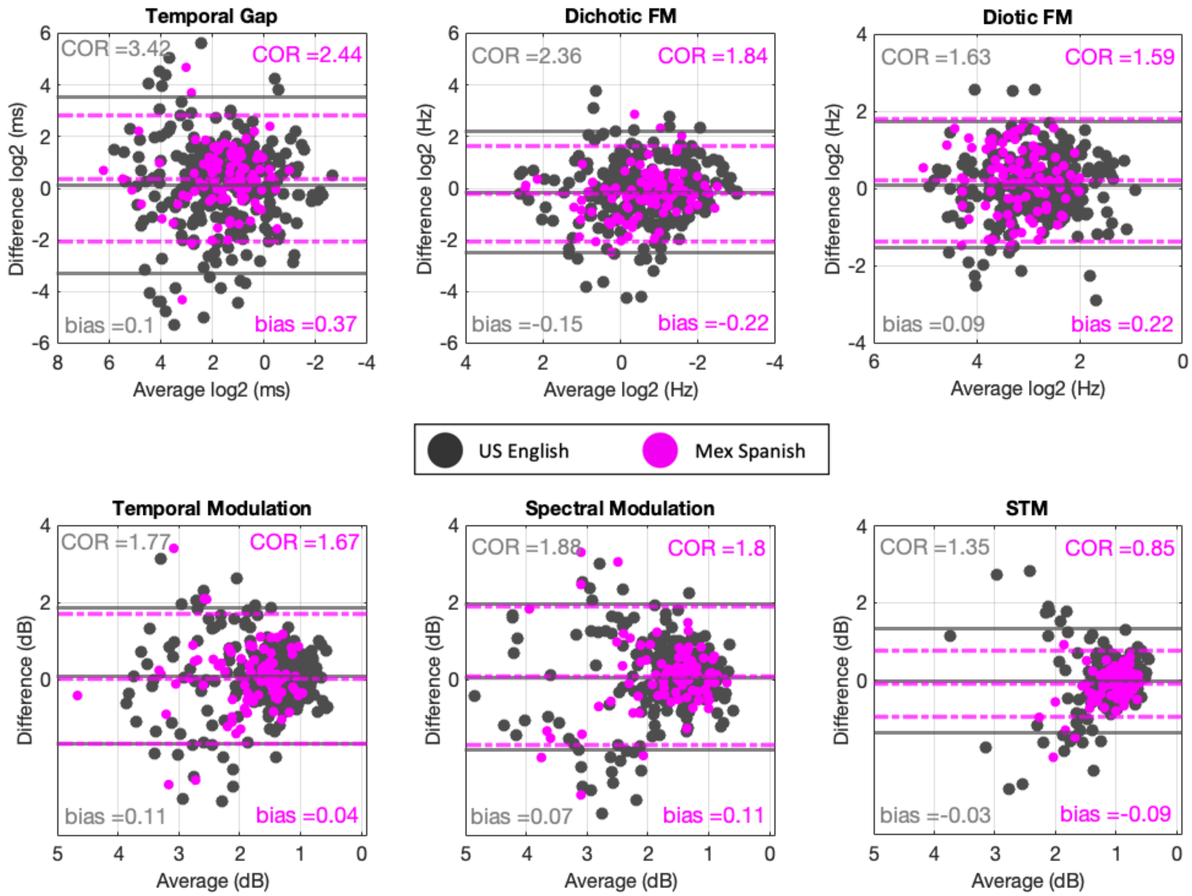

Figure S3. Limits of agreement between sessions for the combined datasets of Larrea-Mancera et al., 2020; 2022 (English in black) and this study (Spanish in pink). Axes were inverted so that better performance is oriented towards the right. The mean difference between sessions and the limits of agreement are indicated with dotted lines and the same color code. A measure of bias, and a coefficient of reliability are also provided for each distribution.

3. Mean comparisons of non-speech measures

A series of independent samples t-tests evaluated whether the estimated thresholds were significantly different across datasets (English vs Spanish). These tests were intended to validate that the young listeners without hearing difficulty were comparable to some extent across studies. We do note that even when participants achieved similar thresholds on the speech tests, there might be language differences between the tests, and that in theory different hearing abilities could actually result in similar scores on the two different tests in the same way that similar hearing abilities could result in different scores on the two tests. This is less likely to occur for the non-speech tests, where similar thresholds are more likely to reflect a similar underlying ability, as the non-speech sounds were the same and only the instructions differed (and we did our best to equate the instructions across languages).

Table S2 shows the independent sample t-tests comparing each estimated threshold across 2 sessions of eight different assessments. Different statistics are provided to facilitate an exploration of the differences found across studies in all measures. An alpha value of .05 corrected for multiple comparisons (16) using the Bonferroni method corresponds to .003, and so this corrected value was used to evaluate the significance of the tests. Mann-Whitney U's are provided because they are robust to the normality assumption violation in comparison to the Student t, and this violation is a common trait of our measures. Further, Bayesian versions of both the Student t and the Mann-Whitney U are provided. Increasing Bayes factor (BF_{10}) values indicate increasing levels of evidence against the null hypothesis of no difference across datasets. Strong evidence against the null is considered above a value of 10. As can be observed in this Table S2, most non-speech tests show no difference across studies/languages in either session. Of note, the few marginally significant cases are inconsistent across different mean comparison metrics, which argues against rejecting the null hypothesis without collecting further data. This is not the case in the Speech tests (especially the SRM separated case) where a more consistent pattern of significant differences can be observed across metrics.

Table S2. Independent samples t-tests comparing the merged datasets of Larrea-Mancera et al., 2020, and 2022 in US-English to the Mexican Spanish dataset collected in this study. Parametric and non-parametric tests are provided as well as their Bayesian versions (see Wagenmakers et al., 2018). Even when most distributions significantly deviate from normality, we have provided parametric tests to ease comparison with previous reports from our group using such statistics. Significance of the mean comparisons is flagged with an asterisk (*) when p values <.003, and in case of at least strong evidence for Bayesian tests.

Assessment	Independent Samples T-Test	Statistic	df	p	Cohen's d	BF_{10}	Test of Normality (Shapiro-Wilk) W	W	p
Gap	Student	-0.716	376	0.474	-0.086	0.168	English dataset	0.987	0.01
	Mann-Whitney	12287.500		0.507	-0.046	0.19	Spanish dataset	0.943	< .00
TM	Student	-2.216	368	0.027	-0.269	1.361	English dataset	0.914	< .00
	Mann-Whitney	9582.000		0.003	-0.205	2.26	Spanish dataset	0.875	< .00
SM	Student	0.387	372	0.699	0.046	0.141	English dataset	0.875	< .00
	Mann-Whitney	12423.000		0.885	-0.010	0.132	Spanish dataset	0.854	< .00
STM	Student	1.248	368	0.213	0.149	0.275	English dataset	0.691	< .00
	Mann-Whitney	12689.500		0.765	0.021	0.137	Spanish dataset	0.675	< .00
DichoticFM	Student	-0.152	373	0.879	-0.018	0.132	English dataset	0.964	< .00
	Mann-Whitney	12249.500		0.483	-0.048	0.131	Spanish dataset	0.954	0.00
DioticFM	Student	-3.594	370	< .001*	-0.429	57.035	English dataset	0.994	0.30
	Mann-Whitney	9691.000		< .001*	-0.244	7.813	Spanish dataset	0.978	0.11
Colocated	Student	3.367	370	< .001*	0.399	27.313	English dataset	0.891	< .00
	Mann-Whitney	17433.500		< .001*	0.316	11.893	Spanish dataset	0.943	< .00
Separated	Student	4.077	374	< .001*	0.483	316.986	English dataset	0.972	< .00
	Mann-Whitney	16507.000		< .001*	0.242	16.291	Spanish dataset	0.969	0.02
Spatial Release	Student	-2.946	370	0.003	-0.348	7.882	English dataset	0.983	0.00
	Mann-Whitney	10451.000		0.001*	-0.222	2.939	Spanish dataset	0.966	0.01
Second sessions									
Gap2	Student	0.857	376	0.392	0.103	0.186	English dataset	0.992	0.16
	Mann-Whitney	14093.500		0.213	0.086	0.128	Spanish dataset	0.942	< .00
TM2	Student	-2.939	368	0.004	-0.358	7.838	English dataset	0.870	< .00
	Mann-Whitney	8199.000		< .001*	-0.295	8.02	Spanish dataset	0.870	< .00
SM2	Student	0.476	372	0.634	0.057	0.147	English dataset	0.840	< .00
	Mann-Whitney	12202.000		0.863	-0.012	0.126	Spanish dataset	0.802	< .00
STM2	Student	0.720	368	0.472	0.087	0.17	English dataset	0.782	< .00
	Mann-Whitney	11657.000		0.866	-0.012	0.13	Spanish dataset	0.808	< .00
DichoticFM2	Student	-0.710	373	0.478	-0.084	0.166	English dataset	0.980	< .00
	Mann-Whitney	12265.500		0.467	-0.050	0.146	Spanish dataset	0.972	0.03
DioticFM2	Student	-2.223	370	0.027	-0.265	1.362	English dataset	0.985	0.00
	Mann-Whitney	10719.000		0.012	-0.173	0.455	Spanish dataset	0.978	0.12
Colocated2	Student	4.092	370	< .001*	0.488	335.21	English dataset	0.917	< .00
	Mann-Whitney	17044.500		< .001*	0.345	81.683	Spanish dataset	0.950	0.00
Separated2	Student	4.394	374	< .001*	0.523	1086.47	English dataset	0.943	< .00
	Mann-Whitney	16050.000		< .001*	0.252	8.456	Spanish dataset	0.943	< .00
Spatial Release2	Student	-2.508	370	0.013	-0.298	2.573	English dataset	0.975	< .00
	Mann-Whitney	10425.500		0.007	-0.187	4.009	Spanish dataset	0.948	< .00

4. HHIA

a. The following instructions were given to participants:

Instrucciones: El propósito de esta escala es identificar los problemas que su pérdida auditiva le puedan estar ocasionando. Marque SI, A VECES, o NO para cada pregunta. No se salte una pregunta si evita esa situación por problemas de escucha. Si utiliza algún aparato auditivo, por favor conteste pensando en cómo escucha sin el aparato.

b. The following 25 questions constitute the instrument:

1.Hay algún problema con su escucha el cual le provoca usar menos el teléfono de lo que le gustaría?
2.Hay algún problema con su escucha el cual le causa vergüenza cuando usted conoce por primera vez a las personas?
3.Hay algún problema con su escucha el cual le provoca evitar grupos de gente?
4.Hay algún problema con su escucha el cual le provoca estar irritable?
5.Hay algún problema con su escucha el cual le causa sentirse frustrado al hablar con miembros de su familia?
6.Hay algún problema con su escucha el cual le provoca dificultades cuando va a fiestas?
7.Hay algún problema con su escucha el cual le provoca dificultad para entender/escuchar a colegas de trabajo o clientes?
8.Se siente usted con una desventaja física causada por un problema para escuchar?
9.Hay algún problema con su escucha el cual le provoca dificultad cuando visita amigos, parientes o vecinos?
10.Hay algún problema con su escucha el cual le provoca frustración al hablar con colegas de trabajo o clientes?
11.Hay algún problema con su escucha el cual le provoca dificultad en el cine o cuando ve una película?
12.Hay algún problema con su escucha el cual le provoca estar nervioso/a?
13.Hay algún problema con su escucha el cual le provoca visitar a amigos, parientes o vecinos menos de lo que le gustaría?
14.Hay algún problema con su escucha el cual le provoca tener discusiones con miembros de la familia?
15.Hay algún problema con su escucha el cual le provoca dificultad cuando oye la televisión o el radio?
16.Hay algún problema con su escucha el cual le provoca ir de compras menos de lo que le gustaría?
17.Hay algún problema o dificultad con su escucha el cual le provoca preocupación?
18.Hay algún problema con su escucha el cual le provoca preferir estar solo?
19.Hay algún problema con su escucha el cual le provoca querer hablar menos con sus familiares?
20.Siente usted que cualquier dificultad con su escucha lo limita o le pone obstáculos en su vida personal o social?
21.Hay algún problema con su escucha el cual le provoca dificultades cuando está en un restorán con parientes o amigos?
22.Hay algún problema con su escucha el cual le provoca sentirse deprimido?
23.Hay algún problema con su escucha el cual le provoca usar la televisión o el radio menos de lo que le gustaría?
24.Hay algún problema con su escucha el cual le provoca sentirse incómodo cuando habla con amigos?
25.Hay algún problema con su escucha el cual le provoca sentirse excluido cuando se encuentra con un grupo de personas?

5. Stimulus spectral properties

In this section we provide a series of spectral analyses of the Spanish and English CRM, intended to provide the reader with visual inspection tools to compare the stimuli used. To perform these analyses, all 256 sentences from all 8 talkers in each corpus were normalized to an RMS value of .1 volts in Matlab 2021a (Mathworks, Natick, MA) and the “fft” function was used to calculate the power spectral density. The output of the fft function was converted into dB SPL with a reference value of 65 dB SPL for a .1 Volt RMS. The panels of Figure S4 show the average power spectral density distributions for the four male and four female talkers in the Spanish and English corpuses. In general, the differences among talkers are as large as the differences between corpuses, with the largest differences occurring between the male and female talkers in the region of 100-200 Hz, reflecting differences in the fundamental frequencies of male and female voices.

The fairly small spectral differences between the English and Spanish speech stimuli seem unlikely to be able to explain the differences in performance observed on the spatial release task. This suggested to us that perhaps prosodic differences are where the two corpuses differ most (apart from the linguistic differences, of course). To examine prosodic differences, a modulation spectrum analysis was conducted, the results of which are shown in Figure S5. Following the methods of Gallun and Souza (2008), the signals were halfwave rectified and lowpass filtered to obtain the envelopes and then zero-padded out to five-second duration to allow an FFT to be calculated with a .2 Hz resolution. The frequency components below 40 Hz are plotted, as this is the region in which modulation is perceived as fluctuation rather than as pitch. To facilitate comparison between the corpuses, the modulation was normalized relative to the DC component in the bottom panels of Figure S5. As with the spectral power, the differences among talkers appear to be greater than the differences between the Spanish and English sentences, but there is an indication of slightly different regularities in the modulation patterns of the sentences from the two languages. Future work examining these differences could be worth conducting but is beyond the scope of this analysis.

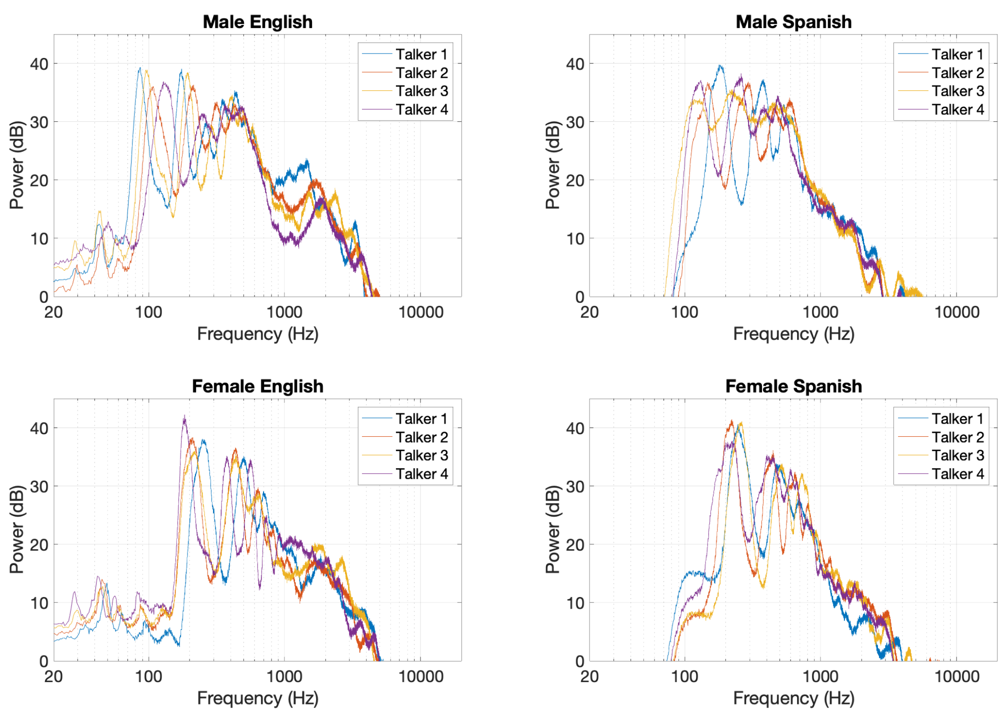

Figure S4. Power spectrum on the top panels and the sound pressure waves in the bottom. The Spanish CRM is represented in the left and the English CRM is represented on the right.

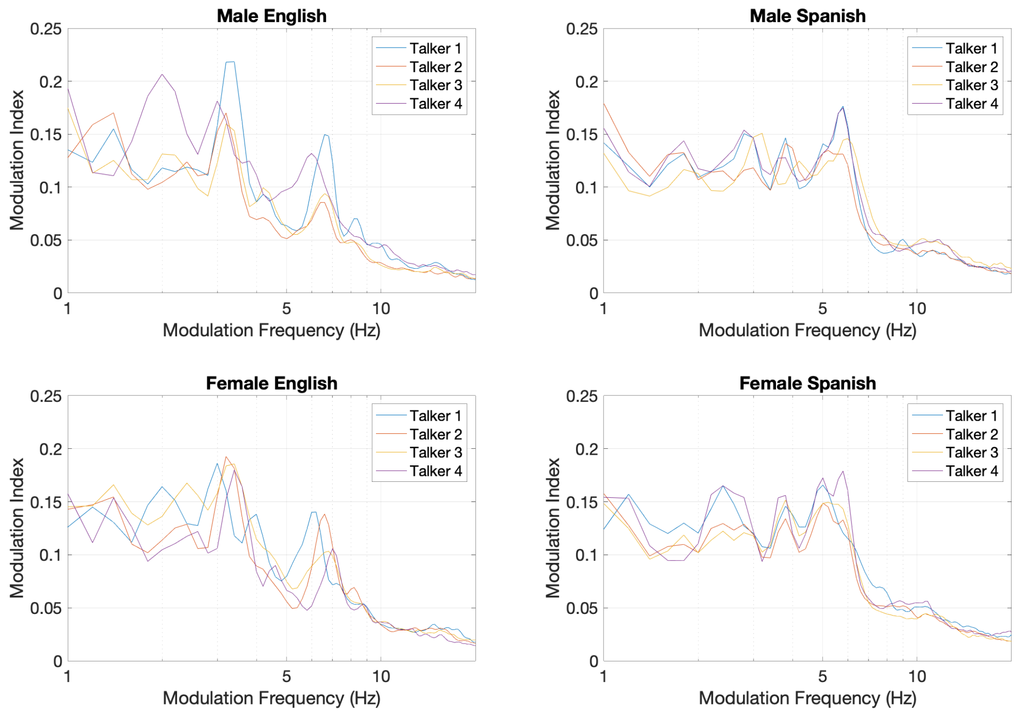

Figure S5. Modulation index (see Souza & Gallun, 2008) for the sentences of the male and female talkers in the Spanish and English Corpuses. Modulation index reflects the modulation power spectral density relative to the 0 Hz component.