Supplementary Material for:

Effects of Formant Proximity and Stimulus Prototypicality on the Neural Discrimination of Vowels: Evidence from the Auditory Frequency-Following Response

T. Christina Zhao[1], Matthew Masapollo[2], Linda Polka[3], Lucie Ménard[4], & Patricia K. Kuhl[1]

[1] Institute for Learning and Brain Sciences, University of Washington, Seattle, WA
[2] Department of Speech, Language and Hearing Sciences, Boston University
[3] School of Communication Sciences and Disorders, McGill University
[4] Department of Linguistics, University of Quebec at Montreal

S.1 Speech Synthesis

The stimuli were synthesized using the Variable Linear Articulatory Model (see Ménard *et al.* (2004). This model is based on a five-formant cascade synthesis system (Feng, 1983). The source was a pulse train generated by the Liljencrants-Fant model (Fant, Liljencrants, & Lin, 1985). The parameters related to the source (glottal symmetry quotient and open quotient) were equal to 0.8 and 0.7, respectively. The value of the first bandwidth (B_1) was calculated according to Fant (1972). So, as a result, there might be differences in harmonic amplitudes, but this is due to the combination of models.

S.2 Acoustic Description of the Speech Stimuli

Prototype	F_1	F_2	F_3	F_4	F_5	B_1	B_2	B_3	B_4	B_5
French	275	790	2522	3410	4159	85	30	35	20	35
English	275	979	2522	3410	4159	85	30	35	20	35

Values of the lower formants (1-5) and their corresponding bandwidths for the less-focal/English /u/ and more-focal/French /u/ prototypes.

S.3 Acoustic and Neural Waveforms for Each Vowel Stimulus

Each vowel stimulus was 100-ms long with a 50-ms onset/offset ramp and had a mean F_0 (or first harmonic, H1) of 130 Hz. Acoustic and neural FFR waveforms for each vowel are shown below in Figure S1.

S.4 Analysis of the Neural Encoding of the Fundamental Frequency (F_0)

To assess the neural encoding of the fundamental frequency (F_0), we performed a 2 × 2 repeated measures analysis of variance (Vowel Type [less/focal/English /u/ vs. more focal/French /u/] × Condition [standard vs. deviant]) on the power (mV2) values in the frequency region corresponding to F_0 (around 135 Hz). The results (shown in Figures 2B [in the manuscript] and S2 [below]) revealed a highly significant main effect of Vowel Type [$F(1,18)=11.538$, $p=0.003$, $\eta^2_p=0.391$], such that there were greater power values observed for the less-focal/English prototypic /u/ (mean=2199.41; 95% CI [1784.33 2614.49]) compared to the more-focal/French prototypic /u/ (mean=1321.32; 95% CI [969.56 1673.08]). The main effect of Condition did not reach statistical significance [$mean_{Standard}=1867.96$, $mean_{Deviant}=1652.77$; $F(1,18)=3.965$, $p=0.062$, $\eta^2_p=0.181$]. There was also no significant interaction [$F(1,18)=0.056$, $p=0.816$, $\eta^2_p=0.003$]. Taken together, these results indicate that there is more robust neural encoding of F_0 for the less-focal but more prototypical exemplar of the /u/ category.

While it may be tempting to conclude that this reflects an enhancement in the neural processing of auditory patterns that are more typical in the native-language (Kuhl & Iverson, 1995; Kuhl *et al*., 2008), these findings are equivocal. Closer inspection of the spectral slices of the stimuli (shown in Figure 2A) indicate that the F_0 values are actually different between the two vowel tokens. In addition, the spectral peak corresponding to F_0 is attenuated in the more-focal/French prototypic /u/ compared to the less-focal/French prototypic /u/. Thus, in the absence of cross-language data from both English- and French-speaking listeners, we cannot draw firm conclusions about whether the enhanced encoding of F_0 for the less-focal/English /u/ prototype is attributable to physical differences in the stimuli and/or long-term linguistic experience.

Supplementary References

Kuhl, P.K. & Iverson, P. (1995). Linguistic experience and the “perceptual magnet effect.” In: Strange, W, (Ed.) *Speech perception and linguistic experience: Issues in cross-language research.* (pp. 121-154). York Press: Baltimore, MD.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M. & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B*, 363, 979-1000.

Ménard, L., Schwartz, J.-L., & Böe, L.-J. (2004). The role of vocal tract morphology in speech development: Perceptual targets and sensori-motor maps for French synthesized vowels from birth to adulthood. *Journal of Speech, Language, and Hearing Research*, 47(5), 1059-1080.

Figures for Supplementary Material

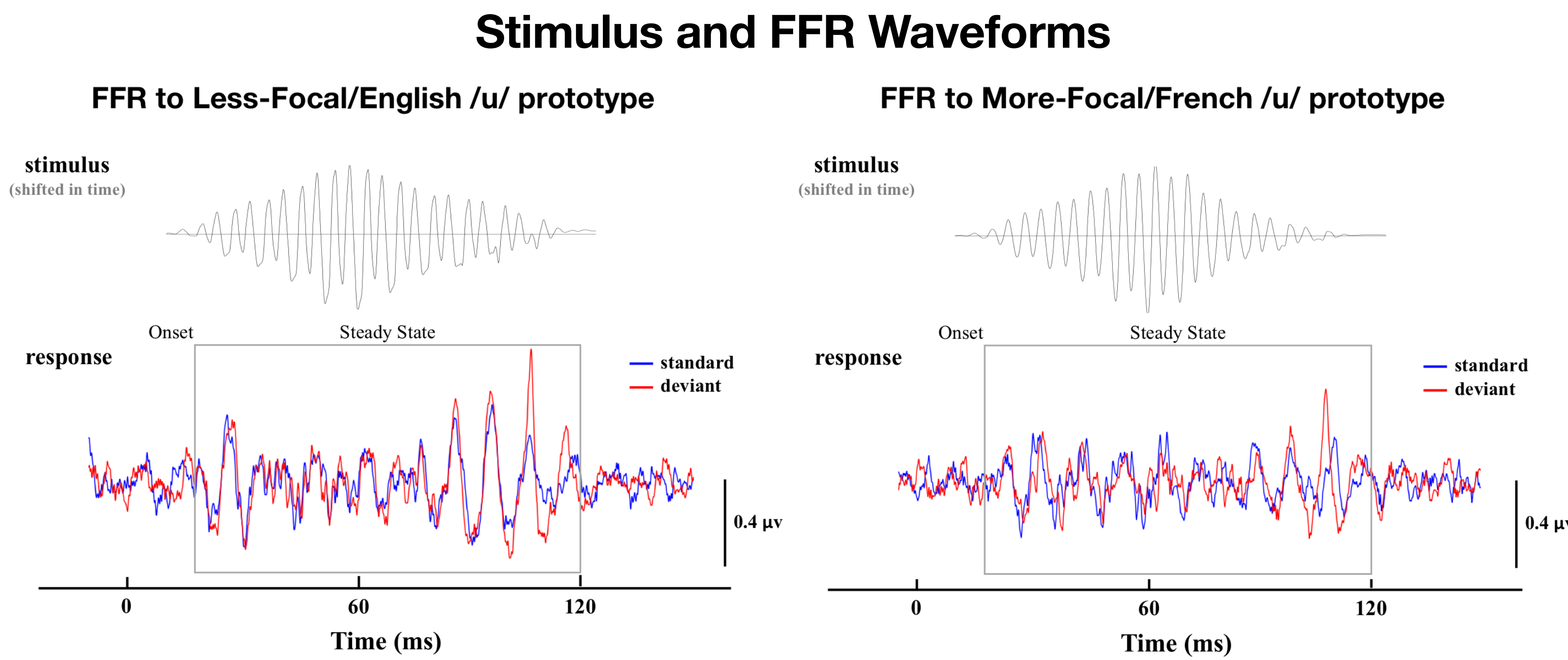

Figure S1: Stimulus (top) and neural response (bottom) waveforms for the two vowel stimuli (less-focal/English /u/ prototype vs. more-focal/French /u/ prototype).

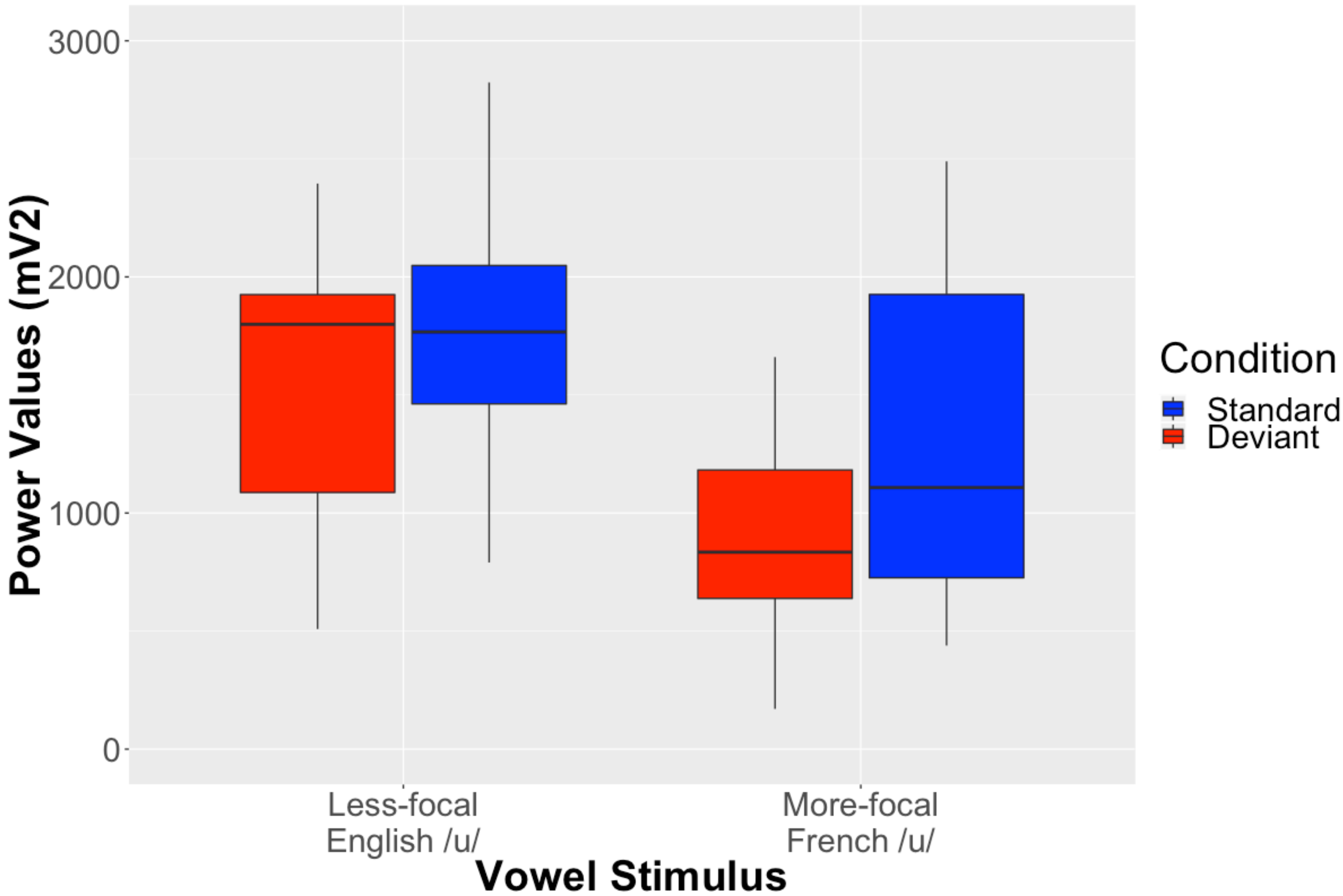

Figure S2: Boxplots of the mean power (mV2) values at the frequency region corresponding to the fundamental frequency (F_0) for each stimulus (less-focal/English /u/ prototype vs. more-focal/French /u/ prototype) as a function of condition (standard vs. deviant).