Supporting Information

**Table S1**

*Acoustic measurements with formant values in Bark scale and F0 values in Hz.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Token | Onset Vowel (ms)  | Syllable Duration (ms) | F1 (Bark) | F2 (Bark) | F3 (Bark) | F0 initial (min-max) (in Hz) | F0 middle (min-max) (in Hz) | F0 final (min-max) (in Hz) |
| sԑ33 1(Standard) | 98 | 583 | 6.44 | 10.14 | 14.44 | 203 (195-218) | 192 (188-195) | 192 (188-196) |
| sԑ33\_2(Standard) | 100 | 584 | 6.55 | 10.57 | 14.94 | 211 (202-222) | 196 (191-202) | 188 (186-192) |
| si33\_1(Vowel Deviant) | 101 | 584 | 3.54 | 12.14 | 15.45 | 217 (207-227) | 193 (188-208) | 187 (183-190) |
| si33\_2(Vowel Deviant) | 97 | 578 | 3.77  | 11.95 | 15.32 | 202 (194-215) | 192 (188-198) | 190 (187-197) |
| sԑ25\_1(Tone Deviant) | 101 | 586 | 6.72 | 9.72 | 14.46 | 208 (199-220) | 196 (193-199) | 230 (199-266) |
| sԑ25\_2(Tone Deviant) | 99 | 579 | 6.73 | 9.70 | 14.44 | 211 (200-222) | 197 (194-202) | 237 (201-268) |

Note. To be able to compare the JND between Standards and Vowel-quality Deviants and Standards and Lexical-tone Deviants we compared the JND in units. For the Vowel-quality Deviant, we transformed the Hz formant values into Bark scale by using the formula: B = (26.81F/ (1960 + F)) – 0.53. We then calculated the Euclidean distance between Standards and Vowel-quality Deviants, which results in a value of 3.39 bark. To obtain the JND units, this value was divided by 0.3 bark, the JND for formant frequencies (Kewley-Port, 2001). The Vowel-quality Deviant spans 11.3 JND units. For the Lexical-tone Deviant, we assumed, according to Jongman et al. (2017), that the JND for rising tones for non-tone listeners is around 6.7 Hz. This results in 10.90 JND at the endpoint of the rising tone. Hence the difference between the Vowel-quality Deviant and Lexical-tone Deviant is about 0.4 JND units suggesting that vowel quality may be more salient than tones. We are aware that the overall (psycho-) acoustic comparison between tones and vowels might not be as simple as we assumed in this calculation.

**Figure S1**

*Oscillogram and spectrogram of the standard*



Note. Oscillogram and spectrogram of /sԑ33/ and the F0 contours of the two different tokens of the syllables /sԑ25/ (red), /si33/ (blue), and /sԑ33/ (green).

Experiment 1

**Table S2.**

*Results from the fitted MMN model with the fixed and random effects.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fixed effects | Estimate (SE) | df | t-value | Pr(>|t|) |
| **Intercept** | **-1.026 (0.233)** | **24.01** | **-4.412** | **<0.001** |
| **Deviant type** | **0.276 (0.057)** | **696.00** | **4.874** | **<0.001** |
| **Region** | **-0.169 (0.057)**  | **696.00** | **2.993** | **<0.01** |
| Deviant type × Region | 0.067 (0.057)  | 696.00 | 1.192 | 0.234 |
| Random effects | Variance | Std. Dev. | Correlation |  |
| Subject (Intercept) | 1.242  | 1.115 |  |  |
| Deviant type | 0.615 | 0.785  | 0.38 |  |

Note. Significant effects are given in bold.

**Table S3.**

*Results from the fitted LDN model with the fixed and random effects.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fixed effects | Estimate (SE) | df | t-value | Pr(>|t|) |
| Intercept | -1.378 (0.268) | 24.00 | -5.141 | <0.001 |
| Deviant type | 0.274 (0.190) | 24.02 | 4.874 | 0.1621 |
| **Region** | **-0.305 (0.056)**  | **672.00** | **-5.441** | **<0.001** |
| **Deviant type × Region** | **-0.113 (0.056)**  | **672.00** | **-2.019** | **0.044** |
| Random effects | Variance | Std. Dev. | Correlation |  |
| Subject (Intercept) | 1.650 | 1.285 |  |  |
| Deviant type | 0.793 | 0.890 | 0.26 |  |

Note. Significant effects are given in bold.

Experiment 2

**Table S4**

*Early analysis time-window*

|  |
| --- |
| Lexical tone deviant |
|  | 6-month-olds | 9-month-olds |
| Region | Estimate (SE) | t-value | p-value | Estimate (SE) | t-value | p-value |
| Frontocentral | 0.312 (0.447) | 0.698 | 1 | **0.890 (0.321)** | **2.775** | **0.03** |
| Posterior | 0.875 (0.440) | 1.987 | 0.23 | -0.039 (0.413) | -0.094  | 1 |
| Vowel-quality deviant |
|  | 6-month-olds | 9-month-olds |
| Region | Estimate (SE) | t-value | p-value | Estimate (SE) | t-value | p-value |
| Frontocentral | **2.349 (0.447)** | **5.261**  | **< 0.001** | 0.268 (0.321) | 0.837  | 1 |
| Posterior | -0.041 (0.440) | -0.094  | 1 | -0.825 (0.417) | -1.978  | 0.21 |

Note. Early analysis time-window: Comparison of the ERP difference waves against zero at frontocentral and posterior regions for both sound contrasts and age groups. Significant MMRs are given in bold.

**Table S5.**

*Results from the fitted early MMR model with the fixed and random effects*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fixed effects | Estimate (SE) | df | t-value | Pr(>|t|) |
| Intercept | 0.474 (0.193)  | 50  | 2.457  | 0.018 |
| Deviant type | 0.036 (0.207)  | 50  | 0.172  | 0.864 |
| **Age**  | **-0.400 (0.193)**  | **50**  | **-2.075**  | **0.043** |
| **Region** | **0.481 (0.105)**  | **1300**  | **4.600**  | **<0.001** |
| Deviant type × Age | 0.316 (0.207)  | 50  | 1.525  | 0.134 |
| **Deviant type × Region** | **-0.390 (0.105)**  | **1300**  | **-3.726**  | **<0.001** |
| Age × Region | 0.024 (0.105)  | 1300  | 0.232  | 0.816 |
| **Deviant type × Age × Region** | **0.349 (0.105)**  | **1300**  | **3.333**  | **<0.001** |
| Random effects | Variance | Std. Dev. | Correlation |  |
| Subject (Intercept) | 1.310  | 1.145 |  |  |
| Deviant type | 1.601  | 1.265  | 0.03 |  |

Note. Significant effects are given in bold.

**Table S6**

*Late analysis time-window*.

|  |
| --- |
| Lexical-tone deviant |
|  | 6-month-olds | 9-month-olds |
| Region | Estimate (SE) | t-value | p-value | Estimate (SE) | t-value | p-value |
| Frontocentral | **3.408 (0.520)** | **6.551** | **<0.001** | **2.497 (0.312)** | **7.997** | **<0.001** |
| Posterior | -0.876 (0.534) | -1.625 | 0.452 | **-1.669 (0.444)** | **-3.758** | **0.001** |
| Vowel-quality deviant |
|  | 6-month-olds | 9-month-olds |
| Region | Estimate (SE) | t-value | p-value | Estimate (SE) | t-value | p-value |
| Frontocentral | 0.518 (0.520) | 0.989 | 1 | **-1.415 (0.312)** | **-3.798** | **0.001** |
| Posterior | **-1.511 (0.534)** | **-2.804** | **0.03** | **-2.076 (0.444)** | **-4.675** | **<0.001** |

Note. Late analysis time-window: Comparison of the ERP difference waves against zero at frontocentral and posterior regions for both sound contrasts and age groups. Significant MMRs are given in bold.

**Table S7.**

*Results from the fitted late MMR model with the fixed and random effects.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Fixed effects | Estimate (SE) | df | t-value | Pr(>|t|) |
| Intercept | -0.218 (0.225)  | 50  | -0.967  | 0.338 |
| **Deviant type** | **0.904 (0.198)**  | **50**  | **4.569**  | **<0.001** |
| **Age**  | **-0.601 (0.225)**  | **50**  | **-2.673**  | **0.010** |
| **Region** | **1.469 (0.112)**  | **1300**  | **13.081** | **<0.001** |
| Deviant type × Age | 0.022 (0.198)  | 50  | 0.111  | 0.912 |
| **Deviant type × Region** | **0.797 (0.112)** | **1300**  | **7.099**  | **<0.001** |
| Age × Region | -0.108 (0.112) | 1300  | -0.964  | 0.335 |
| **Deviant type × Age × Region** | **0.233 (0.112)** | **1300**  | **2.072**  | **0.039** |
| Random effects | Variance | Std. Dev. | correlation |  |
| Subject (Intercept) | 1.904 | 1.380 |  |  |
| Deviant type | 1.327 | 1.152  | -0.10 |  |

Note. Significant effects are given in bold.

**References**

Kewley-Port, D. (2001). Vowel formant discrimination II: Effects of stimulus uncertainty, consonantal context, and training. *The Journal of the Acoustical Society of America*, *110*(4), 2141–2155.

Jongman, A., Qin, Z., Zhang, J., & Sereno, J. A. (2017). Just noticeable differences for pitch direction, height, and slope for Mandarin and English listeners. *The Journal of the Acoustical Society of America*, *142*(2), EL163–EL169. https://doi.org/10.1121/1.4995526