**Supplementary materials**

1. **Acoustical measurements of the stimuli**

Table 1. Mean (SD, range) f0, F1, F2, F3 and duration of vowels /I/ in [ɣIp] and /i/ in [ɣip], and results of the t tests for comparison of each of these measurements.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Multiple-speaker condition | | |
|  | /i/ in giep | /I/ in gip | t (11) |
| f0 max (Hz) | 269(34, 219-321) | 243(60, 126-321) | .89, p = .39 |
| f0 min (Hz) | 226(32, 184-272) | 216(55, 120-294) | .35, p = .56 |
| F1 (Hz) | 320(29, 279-376) | 502(30, 454-556) | **13.72, p <.001** |
| F2 (Hz) | 2651(342, 1818, 3239) | 2636(139, 1996-2465) | **3.75, p = .003** |
| F3 (Hz) | 3189(231, 2724-3624) | 2896(191, 2601-3193) | **3.19, p =.009** |
| Vowel duration (ms) | 91(16, 66-122) | 88(11, 70-113) | .87, p = .41 |
| Initial consonant duration (ms) | 87(2, 83-89) | 88(2, 86-91) | 1.76, p = .12 |
| Word duration (ms) | 346(9, 331-361) | 342(10, 323-357) | 1.04, p=.32 |



Figure 1. Scatterplot of the F1 and F2 values of the vowels in the stimuli.

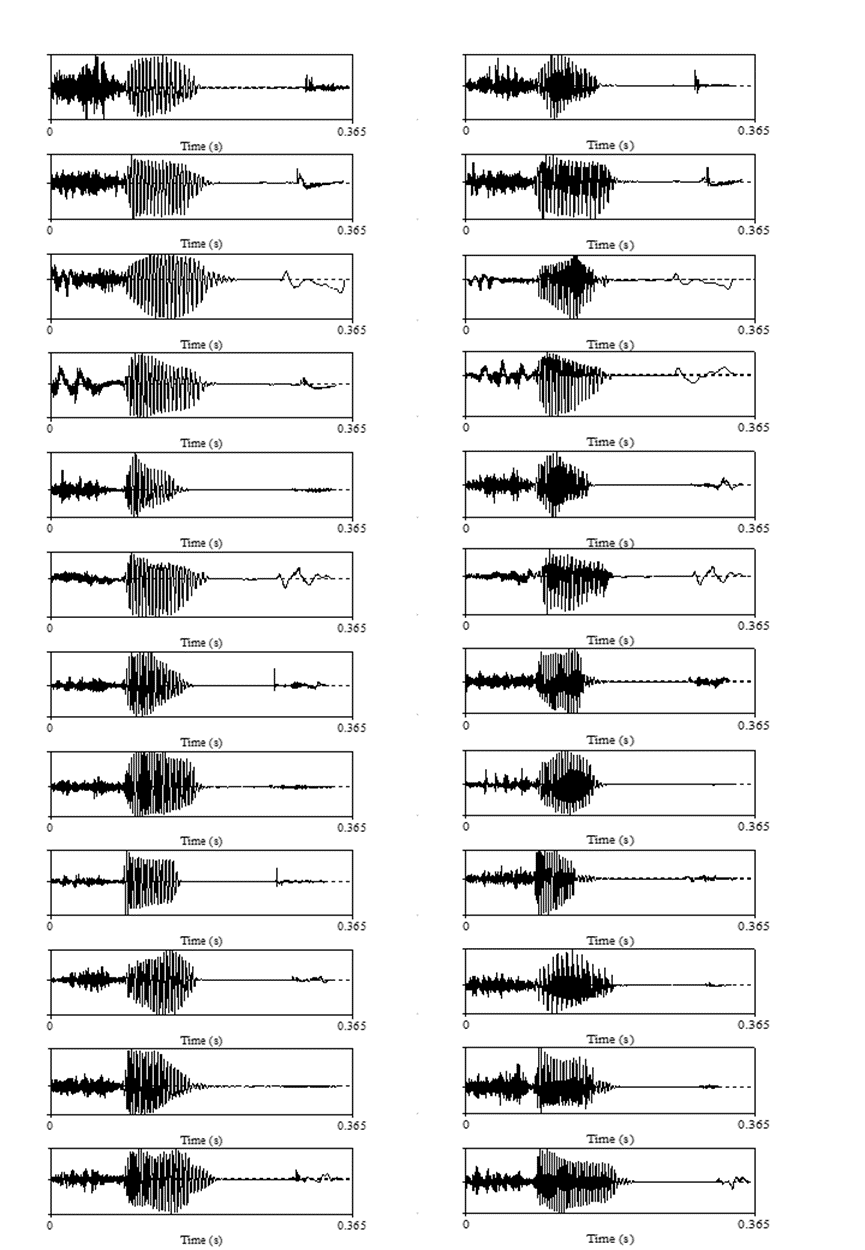


Figure 2. the oscillograms of the 12 gieps (left column) and 12 gips (right column) in the multiple-speaker condition. The 5th pair are from the top are the two tokens used in the single-speaker condition.

1. **Mean peak amplitudes at F3, Fz, F4, C3, Cz, C4 in the multiple- and single-speaker condition, all the participants collapsed.**

Table 2. Mean (SD) amplitude (μv) of the standard ERP (STD), the deviant ERP (DEV), and the p-MMR in the multiple-speaker (MS) and single-speaker (SS) condition at corresponding p-MMR peak latency.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | F3 | Fz | F4 | C3 | Cz | C4 |
| MS | STD | 1.00 (1.81) | 1.10 (1.80) | 1.04 (1.71) | -0.04 (1.42) | -0.12  (1.50) | -0.09  (1.34) |
| DEV | 2.31 (2.82) | 1.89 (2.71) | 1.46 (2.76) | 0.34 (2.73) | -1.42 (3.63) | -0.98 (2.41) |
| p-MMR | 1.32 (2.57) | 0.79 (3.02) | 0.42 (3.18) | 0.39 (2.77) | -1.29 (3.73) | -.089 (2.54) |
| SS | STD | 0.19 (2.03) | -0.20 (1.73) | -0.09 (1.78) | 0.40 (1.65) | 0.03 (1.66) | 0.03  (1.58) |
| DEV | 2.12 (3.47) | 0.64 (3.45) | 0.96 (3.49) | 0.87 (3.47) | -0.97 (3.34) | -0.34 (3.23) |
| p-MMR | 1.93 (3.52) | 0.83 (3.84) | 1.05 (4.00) | 0.47 (3.18) | 1.00 (3.37) | -0.37 (3.72) |

1. **Mean peak amplitude at F3, Fz, F4, C3, Cz, C4 in the multiple- and single-speaker condition, separated by vocabulary level.**

Table 3a. Low comprehenders’ (LC) mean (SD) amplitude (μv) of the standard ERP (STD), the deviant ERP (DEV), and the p-MMR in the multiple-speaker (MS) and single-speaker (SS) condition at corresponding individual p-MMR peak latencies.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | F3 | Fz | F4 | C3 | Cz | C4 |
| MS | STD | 0.82 (1.93) | 0.74 (1.85) | 1.12 (1.83) | -0.37 (1.74) | -0.51 (1.73) | -0.29 (1.44) |
| DEV | 1.85 (3.04) | 0.83 (2.51) | 0.92 (2.67) | -0.51 (2.69) | -2.07 (3.88) | -1.23 (2.69) |
| p-MMR | 1.04 (2.56) | 0.10 (3.08) | -0.20 (3.06) | -0.14 (3.12) | -1.56 (4.59) | -0.94 (2.72) |
| SS | STD | -0.38 (1.98) | -0.41 (1.77) | 0.04 (1.79) | -0.51 (1.59) | 0.03 (1.57) | 0.19  (1.73) |
| DEV | 2.15 (2.93) | 0.24 (3.64) | 0.03 (3.43) | 1.01 (3.32) | -0.34 (3.71) | -0.58 (3.12) |
| p-MMR | 2.53 (3.15) | 0.65 (4.15) | 0.00 (4.13) | 1.06 (3.40) | -0.37 (3.71) | -0.77 (3.98) |

Table 3b. High comprehenders’ (HC) mean (SD) amplitude (μv) of the standard ERP (STD), the deviant ERP (DEV), and the p-MMR in the multiple-speaker (MS) and single-speaker (SS) condition at corresponding individual p-MMR peak latencies.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | F3 | Fz | F4 | C3 | Cz | C4 |
| MS | STD | 1.18 (1.70) | 1.48 (1.70) | 0.96 (1.61) | 0.30 (0.90) | 0.29 (1.10) | 0.11  (1.22) |
| DEV | 2.79 (2.55) | 2.99 (2.51) | 2.03 (2.80) | 1.23 (2.53) | -0.73 (3.28) | -0.72 (2.09) |
| p-MMR | 1.61 (2.66) | 1.52 (2.89) | 1.07 (3.29) | 0.93 (2.36) | -1.02 (2.71) | -0.83 (2.43) |
| SS | STD | 0.79 (1.94) | 0.03 (1.68) | -0.22 (1.81) | 0.87 (1.60) | 0.04 (1.78) | -0.14 (1.43) |
| DEV | 2.09 (4.01) | 1.05 (3.27) | 1.93 (3.35) | 0.71 (3.67) | 1.62 (2.82) | -.10  (3.39) |
| p-MMR | 1.31 (3.82) | 1.02 (3.57) | 2.15 (3.61) | -0.15 (2.86) | -1.66 (2.86) | 0.04  (3.47) |

1. **Standard and deviant ERPs of the high and low comprehenders.**

****

****

Figure 3. The standard [ɣip] and deviant [ɣIp] ERPs of the high (HC) and low comprehenders （LC）in the multiple- (MS) and single-speaker (SS) conditions. Shaded areas represent standard errors.

1. **Productive vocabulary group analysis**

The high producers produced a mean of 221 words (SD = 95), and the low producers produced a mean of 48 words (SD = 30). Grand average ERPs were obtained for the high producers and low producers separately. For each group and for each condition, as for the comprehension groups, the individual MMR peak latencies were identified in the 100 ms window surrounding the corresponding grand average peaks, and the peak MMR amplitude was calculated as the mean amplitude of the 40 ms window surrounding individual average peaks. Table 4 lists the time windows where the standard and deviant ERPs differed significantly and the mean of individual p-MMR peak latencies for the high producer (HP) and low producer (LP) group separately. Similar to the comprehensive groups, grand average p-MMR peak latencies were identified between 300ms and 400ms after the stimulus onset at F3. The LP group had a grand average peak latency of 336 ms and 380 ms and the HC group 336 ms and 352 ms in the multiple- and single-speaker condition respectively. Table 5 shows the mean p-MMR amplitude values.

Table 4. The time windows (ms, after stimulus onset) where the standard and deviant ERP differed significantly and mean p-MMR peak latencies (ms, after stimulus onset) in the multiple-speaker (MS) and single-speaker (SS) conditions, separated for HP (high producers) and LP (low producers). (+) indicates that the deviant ERP was more positive than the standard ERP, and (-) indicates that the deviant ERP was more negative than the standard ERP, and -- indicates lack of significant difference.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| condition | Vocabulary level | F3 | Fz | F4 | Mean (SD) of individual p-MMR peak latencies at F3 |
| MS | HP | 236-256(+)  296-384(+) | 176-192(+)  324-348(+) | 368-380(+) | 344 (27) |
|  | LP | 328-352(+) | -- | 544-576(-) | 335 (31) |
| SS | HP | -- | -- | -- | 350 (32) |
|  | LP | 160-264(+)  292-472(+) | -- | -- | 375 (35) |

Table 5. Mean (SD) p-MMR amplitude at F3, Fz, F4, and C3, Cz, C4 for the multiple-speaker (MS) and single-speaker (SS) conditions, separated by vocabulary level. HP = high producers, and LP = low producers.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| condition | Vocabulary level | F3 | Fz | F4 | C3 | Cz | C4 |
| MS | LP | 0.98 (2.56) | 0.49 (3.24) | -0.13 (3.38) | 0.34 (3.12) | -2.19 (4.34) | -0.95 (2.48) |
|  | HP | 1.66 (2.64) | 1.11 (2.86) | 1.00 (2.97) | 0.43 (2.48) | -0.36 (2.84) | -0.82 (2.69) |
| SS | LP | 2.73 (3.10) | 0.88 (3.85) | 0.54 (4.30) | 1.43 (3.51) | -0.95 (3.58) | -0.49 (4.66) |
|  | HP | 1.09 (3.78) | 0.79 (3.91) | 1.59 (3.66) | -0.54 (2.46) | -1.05 (3.21) | -0.25 (2.48) |

The electrode (F3, Fz, F4) \* condition (multiple- or single-speaker) \* production groups mixed effect ANOVA found significant main effect of electrode, F(2, 102) = 5.71, p = .004, partial η2 = .10. The main effect of condition was not significant, F(1, 51) = .58, p = .45, partial η2 = .01, and neither was the effect of production group F(1, 51) = 0.26, p = .61, partial η2 = .01. The interaction between electrode and production group was significant, F(2, 102) = 3.79, p = .03, partial η2 = .07. Bonferroni corrected posthoc pairwise comparison found that for the low producers, their mismatch response at F3 was larger (i.e., more positive) than at Fz or F4, both p < .05, while no significant cross-electrode difference was observed for the high producers.

Therefore, the production and comprehension median split analysis found similar results, namely that first, the MMR amplitudes were comparable across the multiple- and single-speaker condition for both the groups, and second, those who had a smaller vocabulary had a more left lateralized p-MMR than those who had a larger vocabulary. Interestingly, when the participants were split by production scores, the point-by-point t tests found no MMR, either positive or negative, in the single token condition among the high producers. One likely explanation for this null result is that the high producers were at the transitional stage, namely the adult MMN might be emerging in some toddlers but not in the other. The low producers, on the contrary, were still positive responders as a whole.