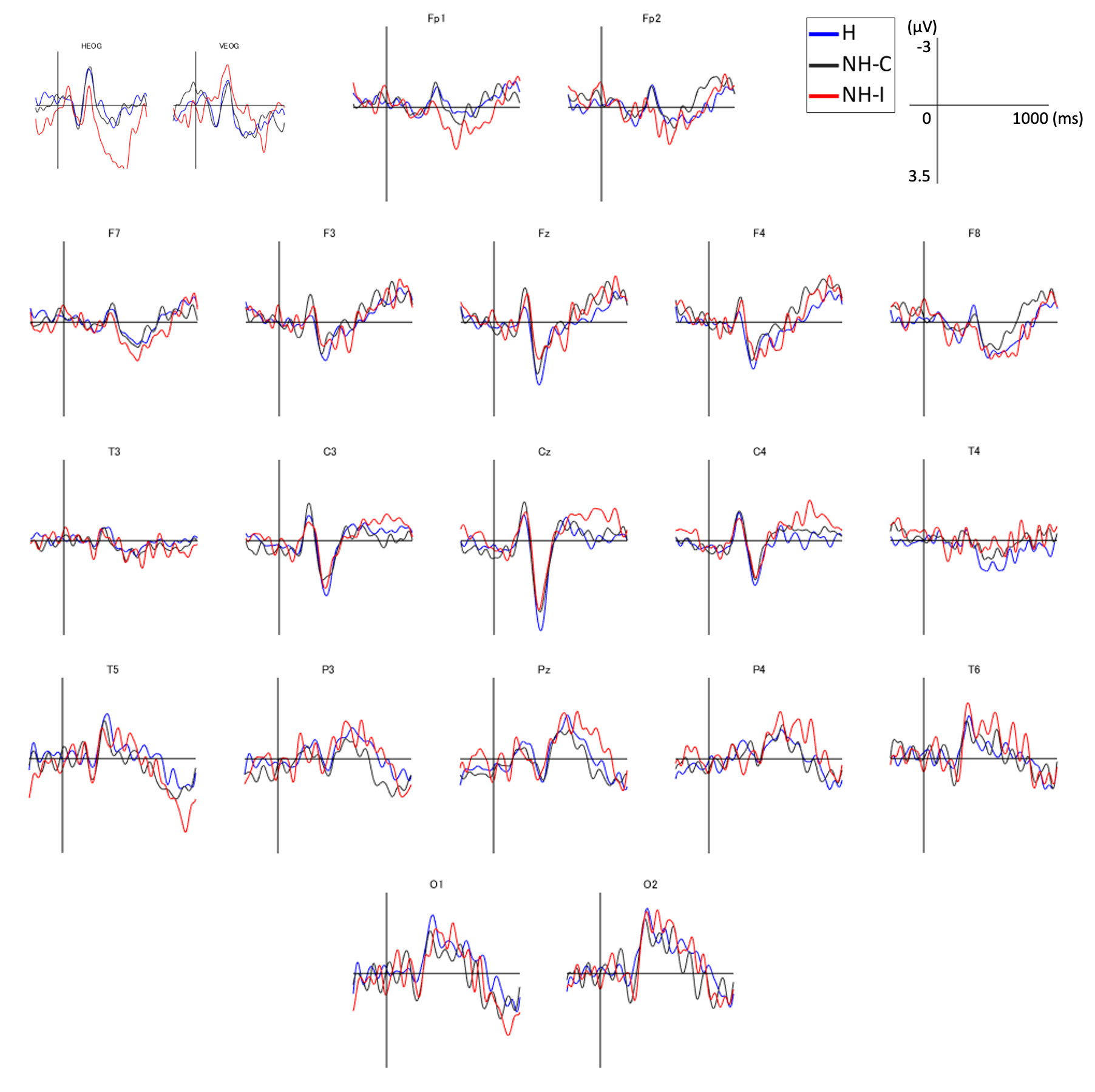
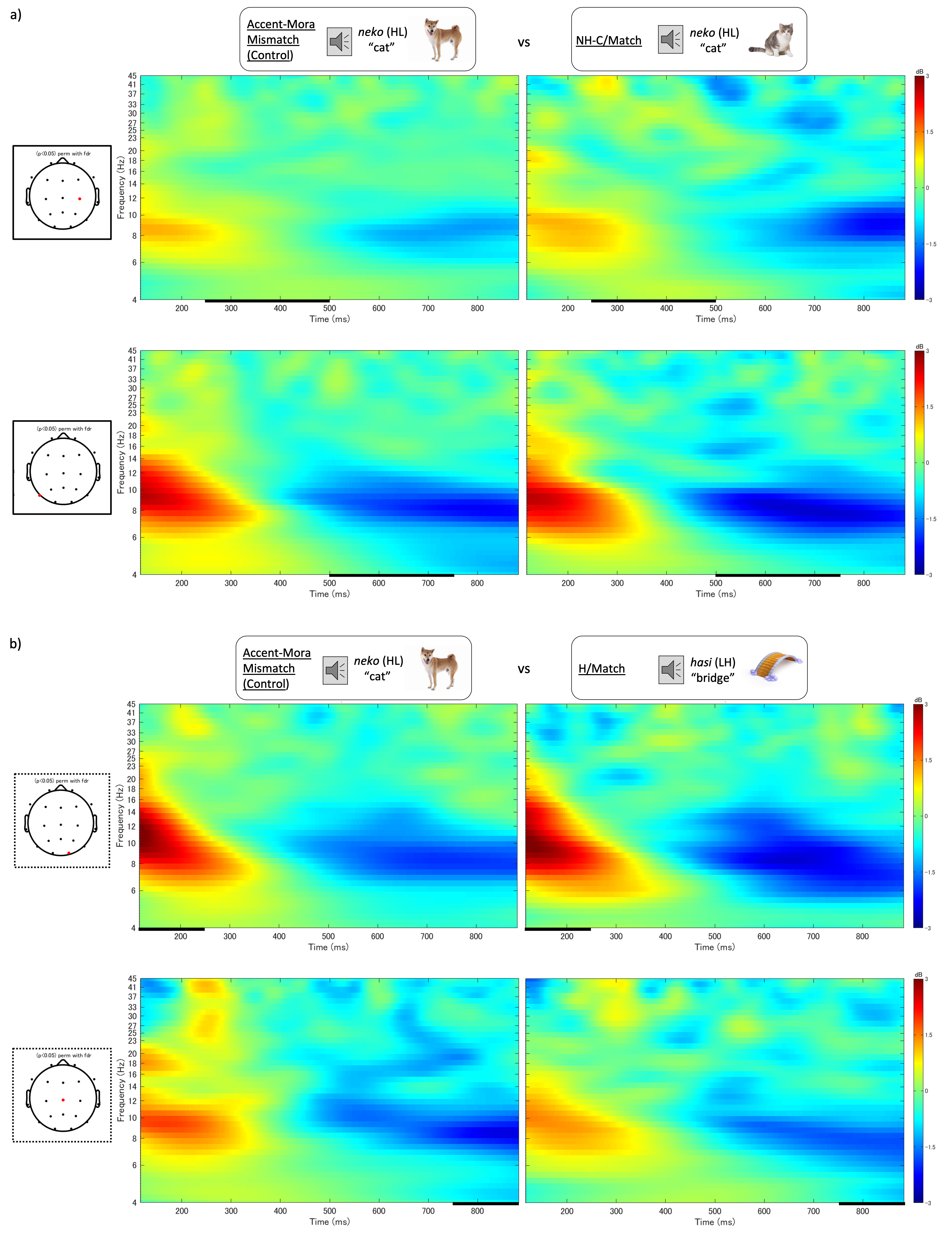
Supplementary Table 1: Mean (SD) acoustic measures of auditory stimuli.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | HL pronunciation  (NH-C for HL words, NH-I for LH words) | | | | LH pronunciation  (NH-I for HL words, NH-C for LH words) | | | |
|  |  | 1st mora | | 2nd mora | | 1st mora | | 2nd mora | |
| Words without Homonyms (NH) | | |  |  |  |  |  |  |  |
| HL words (e.g., *neko*, “cat”) | | |  |  |  |  |  |  |  |
|  | Pitch | 177.2 | (24.2) | 108.6 | (10.6) | 112.3 | (6.3) | 157.4 | (14) |
|  | Intensity | 59 | (3.2) | 53.8 | (2.7) | 55 | (2.6) | 57.3 | (2.7) |
|  | Duration | 208.6 | (40.1) | 225.7 | (51.9) | 189.1 | (39.6) | 243.4 | (41.1) |
| LH words (e.g., *inu*, “dog”) | | |  |  |  |  |  |  |  |
|  | Pitch | 188.4 | (12.8) | 116.9 | (5.6) | 119 | (6.4) | 150.6 | (11.6) |
|  | Intensity | 59.2 | (2.5) | 55.3 | (2.3) | 56.2 | (2.2) | 58.5 | (3) |
|  | Duration | 213.8 | (51.5) | 254.6 | (65) | 212.7 | (47.2) | 240.9 | (56.1) |
| Homonyms (H) | | | | |  |  |  |  |  |
| (e.g., *hasi*, “chopsticks” for HL, “bridge” for LH) | | | | |  |  |  |  |  |
|  | Pitch | 175.1 | (9.5) | 109.5 | (13.7) | 131.2 | (44.9) | 149.3 | (13.4) |
|  | Intensity | 62.1 | (3.8) | 56.2 | (3.1) | 58.9 | (3.3) | 60.38 | (3.3) |
|  | Duration | 181.9 | (39.3) | 245.7 | (38.3) | 191.5 | (36.8) | 254.3 | (55.1) |

Note: HL, high–low pitch accent. LH, low–high pitch accent.



Supplementary Figure 1: ERPs recorded from all electrodes, time-locked to the onset of auditory stimuli. The plotted waveforms were low-pass-filtered at 20 Hz. NH-C, correctly accented words that have no homonymous counterparts. NH-I, incorrectly accented words that have no homonymous counterparts. H, homonyms distinguished by meaning through pitch accents.



Supplementary Figure 2: Time-frequency diagrams of electrodes showing significant differences in the low-gamma band (34–37 Hz) between control and target conditions. The electrodes in red indicate significant differences determined by nonparametric permutation tests (p < .05, FDR corrected). The thick black line indicates the time window during which significant differences were observed in the statistical tests. NH-C, correctly accented words that have no homonymous counterparts. H, homonyms distinguished by meaning through pitch accents. Match, the accent of the preceding auditory stimulus matched the word describing the target visual image. Control, the morae and accent information of the preceding auditory stimulus did not match the word describing the target visual image.