



OPEN

SUBJECT AREAS:
HUMAN BEHAVIOUR
PERCEPTION

Received
12 May 2014

Accepted
2 September 2014

Published
17 October 2014

Correspondence and
requests for materials
should be addressed to
Y.T. (tanaka.yukari.
62x@st.kyoto-u.ac.jp)

Mothers' multimodal information processing is modulated by multimodal interactions with their infants

Yukari Tanaka^{1,5}, Hirokata Fukushima², Kazuo Okanoya^{3,4} & Masako Myowa-Yamakoshi^{1,3}

¹Graduate school of Education, Kyoto University, Kyoto, Japan, ²Faculty of Sociology, Kansai University, Suita, Osaka, Japan, ³Japan Science and Technology Agency, Kawaguchi, Saitama, Japan, ⁴Graduate School of Arts and Science, University of Tokyo, Meguro, Tokyo, Japan, ⁵Japan Society for the Promotion of Science, Tokyo, Japan.

Social learning in infancy is known to be facilitated by multimodal (e.g., visual, tactile, and verbal) cues provided by caregivers. In parallel with infants' development, recent research has revealed that maternal neural activity is altered through interaction with infants, for instance, to be sensitive to infant-directed speech (IDS). The present study investigated the effect of mother-infant multimodal interaction on maternal neural activity. Event-related potentials (ERPs) of mothers were compared to non-mothers during perception of tactile-related words primed by tactile cues. Only mothers showed ERP modulation when tactile cues were incongruent with the subsequent words, and only when the words were delivered with IDS prosody. Furthermore, the frequency of mothers' use of those words was correlated with the magnitude of ERP differentiation between congruent and incongruent stimuli presentations. These results suggest that mother-infant daily interactions enhance multimodal integration of the maternal brain in parenting contexts.

Human caregivers modify their behaviors when interacting with their infants. One typical modification of adults' interaction style is infant-directed speech (IDS), which is characterized by the features of specific prosodic patterns such as higher pitch, greater pitch variations, longer pauses, and a more rhythmic, slower tempo when compared to adult-directed speech (ADS)^{1–3}. IDS has the function of drawing infants' attention^{4–6}, promoting emotional interaction⁷, and facilitating language acquisition in infancy^{8–11}. Importantly, the behavioral modification of caregivers is multimodal in nature, including visual, auditory, and tactile information¹². For example, mothers use pointing gestures to a target object, coupled with IDS¹³. These multimodal cues are often demonstrated with redundant temporal synchrony among several modalities (i.e., multimodal motherese), which emphasizes the salient scenes in the environment, and facilitates infants' learning^{14–16}. Multimodal interaction plays the important role of making ambient references and intentions clear, and encouraging smooth interaction between mothers and infants.

Although parents have been shown to modify their behaviors, it is still unclear how such experiences affect the cognitive and neural functions of caregivers. Recent neuroscience research has revealed that parenting experiences alter the brain activity of mothers. For example, mothers with preverbal infants showed enhanced cortical activation in the auditory dorsal pathway of the language areas (Broca's and Wernicke's areas) during the perception of IDS, whereas fathers and non-parents did not show enhanced activity in these areas¹⁷. These results suggest that daily experiences of vocalization and hearing speech feedback of IDS enhanced mothers' brain activities in these areas, which are considered to reflect the processing of phonological information. However, the previous research focused only on the phonological aspects of IDS perception, and we do not know whether multimodal information processing is enhanced or modulated by parenting experiences.

The aim of the present study is to reveal whether and how mothers' interactions with infants affect mothers' neural processing of multimodal information in the context of parenting behaviors. Specifically, we focused on auditory and tactile multimodal information processing. Auditory information, especially verbal cues, are quite important for word learning in infancy. Auditory cues enable mothers to convey referential information, which directs infants' attention to specific objects or to specific aspects of the environment^{18,19}. Furthermore, tactile cues (e.g., hugging, touching, and kissing) are important signals of interaction between mothers and infants²⁰. Tactile cues are often combined with verbal cues. For example, mothers often use baby talk including mimetic words as infant-directed speech²¹, and provide their toddlers (aged around 2.5 years) with tactile experiences accompanied



by tactile-related onomatopoeia words in an IDS manner²² (e.g., having infants touch a soft blanket and saying “fuwa-fuwa”—Japanese onomatopoeia referring to something soft—in a high-pitched voice).

In order to investigate the integration process of tactile-auditory information, we applied a multimodal semantic priming paradigm^{23,24} (Figure 1). The paradigm consisted of priming a tactile stimulus (prime) followed by an auditory linguistic stimulus (target). Participants were required to respond by choosing the word identical to the target stimuli. These two kinds of stimuli (tactile prime and auditory target) were semantically either congruent or incongruent. Target stimuli occurred in one of two prosodic conditions: IDS and ADS. We measured event-related potentials (ERPs) in mothers and non-mothers and analyzed the data collected during the presentation of the target stimuli. The ERPs of mothers and non-mothers were compared between the conditions of congruency (congruent or incongruent) and prosody (IDS or ADS).

We predicted that mothers' ERPs would show enhanced sensitivity to the congruency of the stimuli, particularly when they were hearing IDS prosody. This is because mothers are assumed to have more experience than non-mothers with multimodal interaction with infants in an infant-directed speech manner. To further support this hypothesis, we also tested the correlation between mothers' ERP responses and their reported frequency of use of tactile-related words during their daily interactions with their infants.

Results

ERP Response of Mothers and Non-mothers. We focused on the results obtained from the middle frontal region (Fz), where the group differences of interest were most evident. ERPs in other regions are described in Figure 2, Table 1, and the supplementary results. ERPs elicited by auditory stimuli showed a negative peak around ~150 ms (N1), a positive peak around ~240 ms (P2), and a negative peak around ~350 ms from the target onset (N400) (See Figure 2 and Table 1). Each component was quantified as the mean amplitude of each of the following periods: N1, 120–180 ms; P2, 200–300 ms; and N400, 300–500 ms from target onset. These amplitudes were analyzed by mixed measures analysis of variance (ANOVAs) with group (2: mothers/non-mothers) as the between-participant factor,

and prosody (2: ADS/IDS) and congruency (2: congruent/incongruent) as within-participant factors.

For the period of N1, we found a significant interaction between group, prosody, and congruency ($F_{1,32} = 4.46$, $p = .04$, $\eta_p^2 = .12$; Fig. 3A and Fig. 3B). We then conducted two-way ANOVAs for each group with prosody and congruency as within-participant factors. We found a significant interaction between prosody and congruency for the mothers group ($F_{1,16} = 6.43$, $p = .02$, $\eta_p^2 = .28$). Mothers showed significant differences between IDS-congruent and IDS-incongruent peaks (mean amplitude for IDS-congruent = $-0.22 \mu\text{V}$, IDS-incongruent = $0.34 \mu\text{V}$, $t = 2.60$, $p = .02$), but not between the ADS prosodies ($t = 0.67$, $p = .51$). Non-mothers did not show significant main effects or interactions ($F_s < 1.05$, $p_s > .32$, *n.s.*). No interactions or main effects were detected in other regions ($F_s < 3.0$, $p_s > .05$, see Supplementary Figure S1).

For the period of P2, we also found significant interactions between group, prosody, and congruency ($F_{1,32} = 4.65$, $p = .04$, $\eta_p^2 = .13$, Fig. 3A and Fig. 3C). Again, two-way ANOVAs for each group revealed a significant interaction between prosody and congruency for only the mothers group ($F_{1,16} = 9.40$, $p < .01$, $\eta_p^2 = .37$). Post-hoc analysis revealed that mothers showed a larger amplitude in the IDS-incongruent condition than the IDS-congruent condition (IDS-congruent = $0.71 \mu\text{V}$, IDS-incongruent = $1.25 \mu\text{V}$, $t = 2.13$, $p = .05$). We also found that only mothers showed a significantly larger amplitude in the IDS-incongruent than the ADS-incongruent condition (ADS-incongruent = $0.47 \mu\text{V}$, IDS-incongruent = $1.25 \mu\text{V}$, $t = 3.57$, $p < .01$). Non-mothers showed no significant main effects or interaction ($p_s > .05$, *n.s.*). Aside from the group-related modulations, a main effect of congruency was observed in the period of P2 in the left frontal regions (F3) with greater positivity in the incongruent condition than in the congruent condition (Fig. 2 and Table 1).

For the period of N400, we did not find a three-way interaction between group, prosody, and congruency ($F_{1,32} = 1.89$, $p = .18$). We found that the interaction between prosody and congruency in the middle central region ($F_{1,32} = 2.80$, $p = .10$, $\eta_p^2 = .09$) was not significant. We also found a main effect of prosody on mean amplitude in the right frontal and central regions, with greater negative mean amplitude in IDS prosody than ADS prosody (Fig. 2 and Table 1).

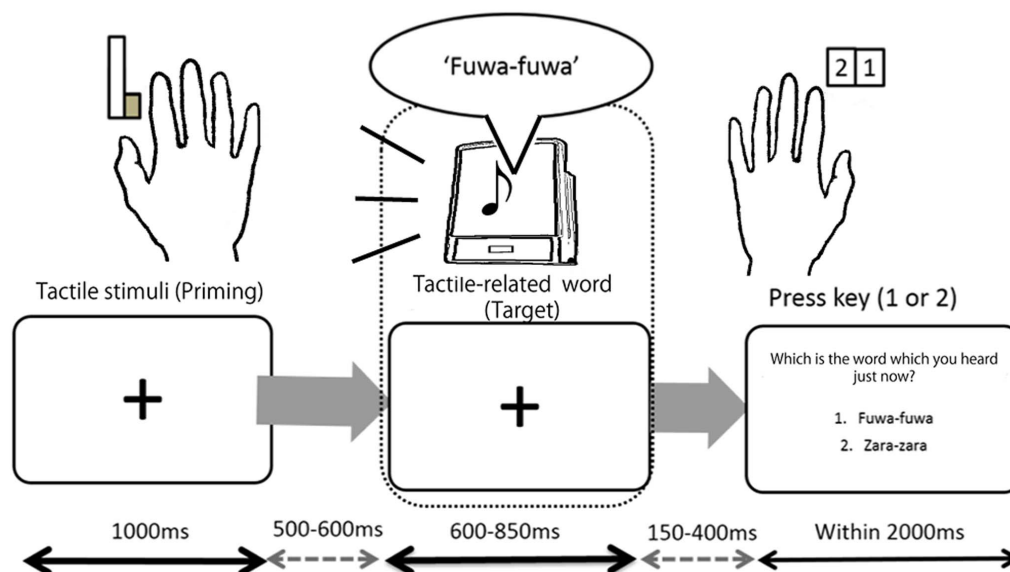


Figure 1 | Experimental procedure of the present study. Participants saw an on-screen fixation point. A tactile stimulus (prime) was then presented followed by a word presentation (target). Participants were required to identify the word they heard by pressing a key. Inter-stimulus intervals between the prime and target were 500 to 600 ms and between the target and identification task were 150 to 400 ms. EEG data recorded during the target stimulus presentation was analyzed.

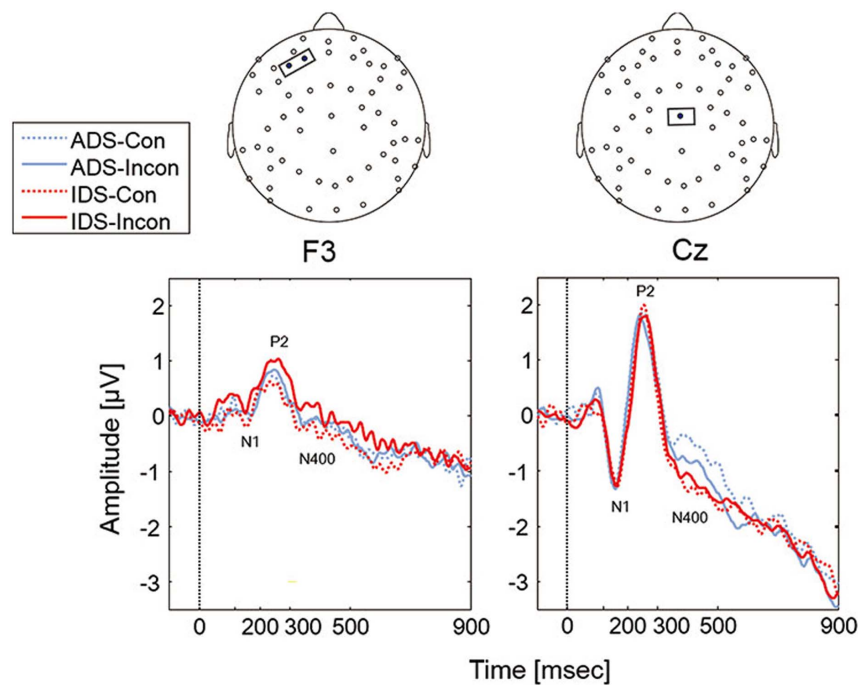


Figure 2 | Grand-averaged ERP waveforms of all participants. Dashed lines designate the congruent condition and solid lines designate the incongruent condition. Line color indicates prosody (blue as ADS or red as IDS). The peak and timeline of each component (N1, P2, and N400) are shown in the ERP waves. ERPs in other regions are shown in the supplementary results (S1).

Relationship Between the ERPs of Mothers and the Frequency of Use of Tactile-related Words. The frequency with which mothers used the target words with their infants in daily interactions was calculated from a parent questionnaire. We conducted correlation analysis between the frequency of mothers' target word use and the effect of audio-tactile congruency on ERPs. The effect of audio-tactile congruency was defined as the ERP differentiation between congruent and incongruent conditions in both ADS and IDS prosodies; mean amplitude in the ADS-incongruent condition minus that in the ADS-congruent condition for ADS prosody, and mean amplitude in the IDS-incongruent condition minus that in the IDS-congruent condition for IDS prosody. We found a significant positive correlation between the frequency of target word usage and the differential amplitude of the P2 component for the IDS condition, but not for the ADS condition (IDS: $r_s = .54$, $p = .03$; ADS: $r_s = -.36$, $p = .15$; Fig. 4). We also found the same pattern of correlation for N400, again only for the IDS prosody (IDS: $r_s = .60$, $p = .01$; ADS: $r_s = -.13$, $p = .61$). The early component (N1) did not provide a significant correlation in either prosody. We found no correlations between the frequency of target word usage and ERP response in any other regions (See Supplementary Table S2).

Discussion

We investigated whether and how maternal multimodal interactions with infants would affect mothers' neural processing of

audio-tactile integration using ERP methodology. We found that only the group of mothers showed differences in ERP amplitudes between the IDS-congruent and IDS-incongruent conditions for N1 latency. This sensitivity to the mismatch between the tactile and verbal stimuli was observed when the tactile-related words were presented with IDS prosody, but not with ADS prosody. Contrary to mothers, the ERPs of the non-mother group did not show the specific sensitivity to the incongruity between the verbal and tactile stimuli with IDS prosody. The auditory N1 component is considered to be associated with the processing of sensory information such as frequency and intensity of a stimulus²⁵. The present finding of N1 modulation after the semantic mismatch of the audio-tactile stimuli is likely to reflect multisensory integration reported with similar latency in a recent ERP study using naturalistic stimuli²⁶.

Furthermore, we found mother-specific ERP responses in middle P2 latency; again only the group of mothers showed the different ERP amplitude between congruent and incongruent stimuli, and only with IDS prosody, whereas the non-mothers group did not show any significant differences between stimuli. The P2 component is assumed to reflect the processing of phonological categorization, being related to the neural representations of multimodal categorization^{27,28}. These results suggest that mothers' modulated ERPs in the IDS condition (i.e., different ERP amplitudes according to the congruency of word and tactile stimuli) are related to the processing

Table 1 | Main Effects of Congruency and Prosody in Fronto-central Regions

Effect	Region	Component	Peak Latency (msec)	Comparison	F	p	η^2
Congruency	F3	P2	247.18	Positive Incon > Con	4.90	.03	.13
	C3	P2	242.00	Positive Incon > Con	6.13	.02	.16
	F4	N400	343.62	Negative IDS > ADS	4.83	.04	.13
	C3	N400	356.15	Negative IDS > ADS	9.86	.04	.24
	Cz	N400	359.24	Negative IDS > ADS	10.36	.003	.25
	C4	N400	358.06	Negative IDS > ADS	15.32	.0001	.32

Note. We calculated the mean amplitude of each component (N1, P2, and N400) in the fronto-central regions, and compared that with group, prosody, and congruency for each region. In the F3 and C3 regions, we found a main effect of congruency in the P2 component, and in the right frontal and central regions, we found a main effect of prosody in the N400 component.

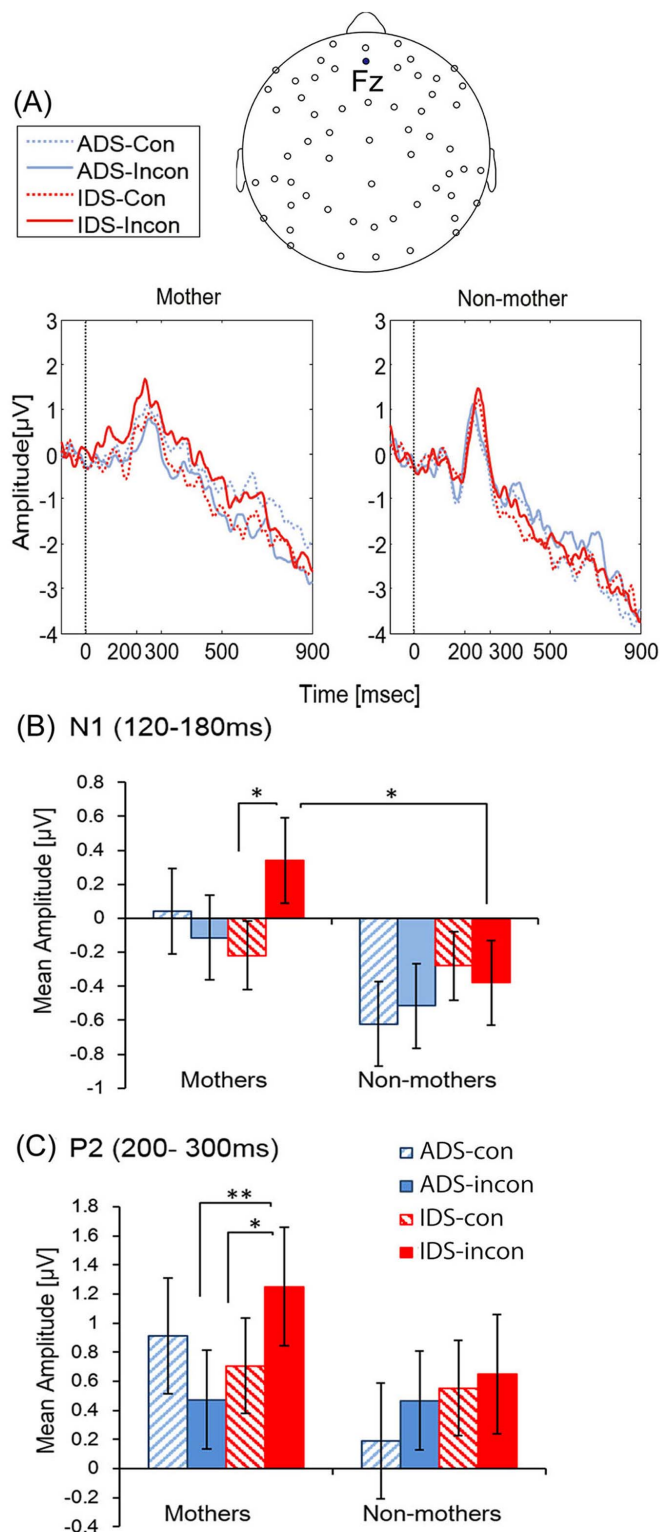


Figure 3 | Grand-averaged ERP waveforms and mean ERP amplitudes of mothers and non-mothers. Grand-averaged ERP waveforms in Fz region (A) are presented for mothers and non-mothers. Mean ERP amplitudes are presented in the period of N1 (B) and P2 (C). Error bars show standard errors. * $p < .05$, ** $p < .01$.

discrimination of multimodal categorization between tactile and verbal cues.

There are some reasons why different ERP modulation between mothers and non-mothers were observed in early-to-middle latency (N1 and P2). One possibility is that mothers are more skilled at

detecting the incongruity of multimodal events than non-mothers. However, both groups showed congruency effects in the left frontal (F3) region, regardless of prosody (Table 1). In other words, the audio-tactile integration process itself was not different between the two groups. Furthermore, mother-specific ERP responses emerged in the IDS prosody condition, not with ADS prosody. These results provide an interesting insight into the functional mechanism of human parenting behavior. Recent studies have suggested that parenting behavior is related to the activation of orbitofrontal regions²⁹, which are considered to have the function of evaluating social rewards and decision-making³⁰. During mother-infant interactions, mothers are required to monitor infants' state and condition, and to respond to and cope with infants' signals quickly. In the present study, participants had to detect and discriminate multimodal congruency. It is possible that verbal cues with IDS prosody motivated mothers to respond selectively to the IDS stimuli, and to evaluate the congruency between tactile and verbal cues, resulting in group differences in early-to-middle latency.

Our data support the hypothesis that one factor influencing mother-specific ERP modulation might be mothers' experiences of speaking tactile-related words to their infants. Correlation analysis revealed that, in the P2 and N400 components elicited from mothers, the difference in amplitudes between the IDS-incongruent and IDS-congruent cues was positively correlated with the frequency of mothers' use of tactile-related words in daily interactions with their toddlers. Again, this effect was observed only in the IDS condition but not in the ADS condition, nor in other regions. We may say that mother-specific ERP responses are an 'experience-dependent effect.' One adult study showed that P2 amplitude is enhanced by the speech training of syllables³¹. The training effect might facilitate mothers' response to IDS stimuli, especially in the incongruent condition, resulting in a larger differential ERP response to IDS prosody.

The middle-to-late component showed relatively clear correlation with the subjective evaluation of the mothers, compared to the single component in the early latency because, in general, the middle-to-late component reflects higher-order conscious processing. In particular, the processing of semantic and category-related information is represented as N400 amplitude, which is an important neurophysiological index for semantic memory organization and conceptual learning³²⁻³⁴. Mothers who often use tactile-related words could have greater accessibility to the semantic meanings of tactile-related words, and they showed larger differential ERPs between IDS-congruent and IDS-incongruent conditions in the late latencies. It is interesting, though, and deserves further investigation that the neural activity of each time scale and the level of multi-modal information processing was related to parenting behavior. It is also important to determine how experience affects neurophysiological responses at different latencies in more detail with other groups, such as childcare workers, grandparents or fathers.

In sum, we found that mothers showed larger ERP responses to IDS-incongruent relative to IDS-congruent stimuli at middle frontal electrodes, whereas non-mothers did not show differential responses. The multimodal congruency effects specific to IDS were related the frequency of using tactile related-words in daily interactions with infants. The results suggest that mother-infant multimodal interaction in daily life enhances mothers' selective neural responses to multimodal information within the parenting context, which might facilitate social cognitive development in infancy.

Methods

Participants. Seventeen mothers (mean age = 32.56 ± 3.76 years, range 25–41 years) parenting toddlers (8 boys, mean age = 20.8 ± 1.65 months, range 19–23 months), and seventeen non-mothers (17 females, mean age = 22.4 ± 2.74 years, range 20–31 years) participated in the study. All participants were neurologically typical, right-handed Japanese speakers, and they were paid for participation. All gave informed consent according to the procedures approved by the Ethics Committee of Web for the Integrated Studies of the Human Mind, Japan (WISH, Japan). Some participants came to the experimental laboratory with their infants. During this time, infants were

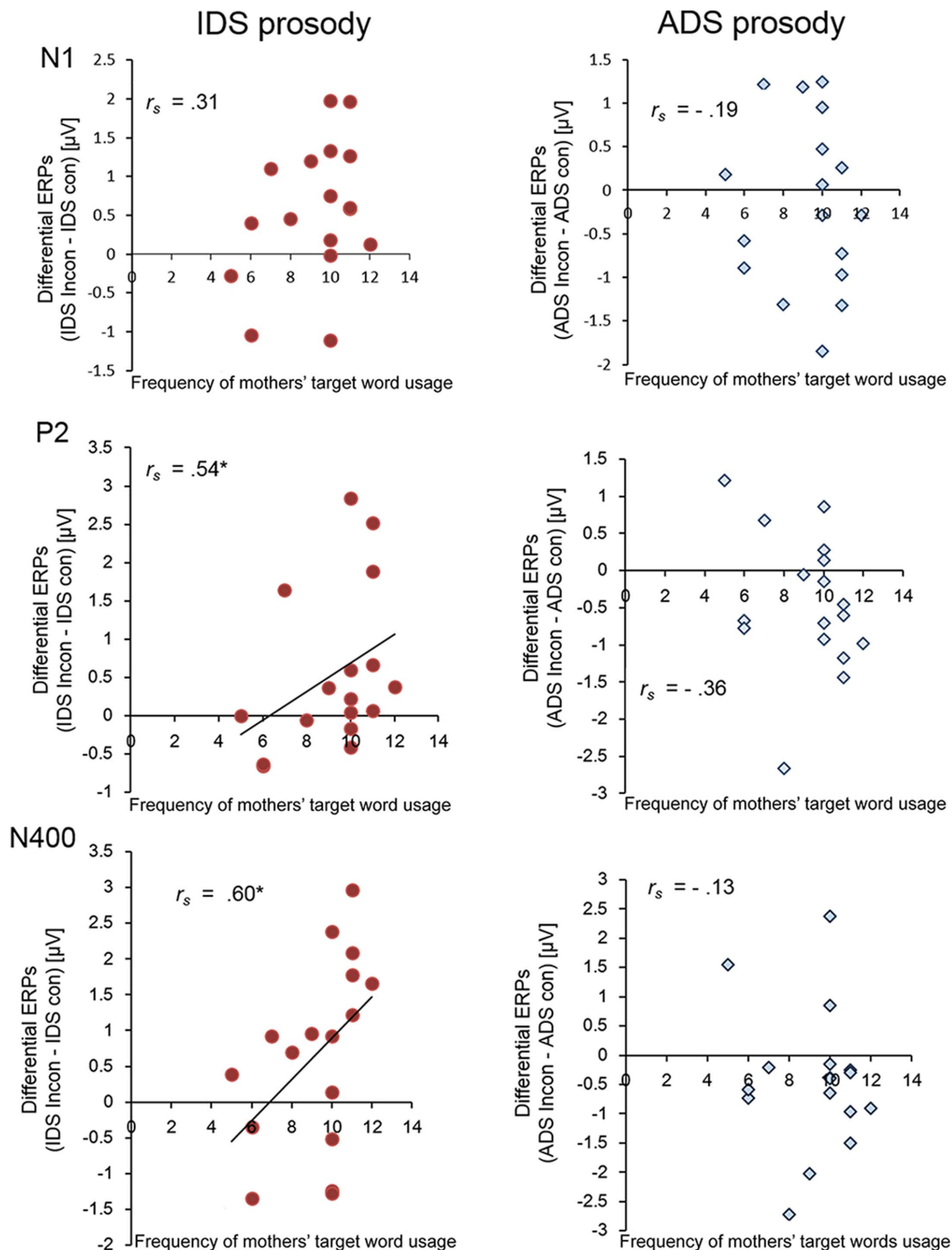


Figure 4 | Correlation between mothers' ERP responses and their usage score for tactile-related words. The graph shows mothers' data. The X-axis shows the frequency of use of tactile-related words and the Y-axis shows differential ERP responses. The graphs represent correlations for N1, P2, and N400. The left side shows differential ERPs with IDS prosody (IDS-incongruent – IDS-congruent) [μ V] and the right side shows those with ADS prosody (ADS-incongruent – ADS-congruent) [μ V].

allowed to explore the room and to play freely in another space with an assistant experimenter. Data from an additional eight mothers and four non-mothers had to be excluded from the subsequent analysis due to muscle artifacts (one mother), extensive eye movement (three mothers), technical problems (three mothers and three non-mothers), infants' crying (one mother), and the lack of participants' attention to the task (one non-mother).

Stimuli. The following three textures were selected as tactile stimuli for the present experiment: fake fur, sand paper, and leather. Each tactile stimulus (length: 3 cm;

width: 2 cm) was attached on a plane surface. Tactile stimuli were placed in a custom-made box so that participants could not see the tactile stimuli (length: 23 cm; width: 31 cm; height: 27 cm). Through an opening in the front (length: 9 cm; width: 11 cm), participants were instructed to put their right hands into the box. They had their right index finger fixed with a band to restrict their body movements (Supplementary Figure S2). The second experimenter sat by the box and presented tactile stimuli manually through an opening at the back of the box (length: 20 cm, width: 27 cm).

As auditory stimuli, we prepared three tactile-related Japanese onomatopoeias as follows: /fuwa-fuwa/ (something soft), /tsuru-tsuru/ (something smooth), and /zara-



zara/ (something rough and hard). These words were selected from a pilot questionnaire given to mothers, which asked how frequently they use tactile-related words with their infants. Final stimuli consisted of three high frequency words in both general and IDS use. The stimuli were recordings of two mothers speaking each word in two prosodic conditions: (i) IDS prosody, in the presence of their toddlers (aged two years) and (ii) ADS prosody, directed at an adult (an experimenter, aged 25 years). The auditory stimuli consisted of a total of 12 stimuli (3 words \times 2 prosodic conditions \times 2 mothers). Words were recorded at a 22.05 kHz sampling rate (in 16-bit monaural) using a digital recorder in a soundproof chamber.

The auditory stimuli were analyzed for the following parameters: average fundamental frequency (F0), pitch maximum (F-Max), frequency range (F-range), and duration. Pitch and duration analyses of the recordings were conducted using Adobe Audition. Statistical analyses were then conducted using Wilcoxon signed-rank test. The analyses, shown in the supplementary information (Supplementary Table S3), indicated that the F0 and F-Max of the auditory stimuli in IDS were significantly higher than those in ADS. F-range of the IDS stimuli was marginally higher than ADS stimuli. Duration was not different between IDS and ADS stimuli because each IDS sample was cut into short single words for use with the ERP paradigm. In order to ensure that IDS stimuli sounded like 'infant-directed speech,' another group of non-parents scored how child-directed (1. Not-at-all (adult-directed) to 7. Very childish) auditory stimuli sounded. IDS stimuli were scored to be more childish than ADS stimuli ($t = 3.38, p = .01$). The intensity of the auditory stimuli was adjusted across stimuli by equalizing the root mean square power of all sound files. These stimuli were presented to the subjects at around 62.50 dB (SPL) sound pressure level.

Procedure. Each trial was comprised of a tactile stimulus (prime) and a subsequent auditory target stimulus (target). The target stimuli were tactile-related words spoken in an IDS (50%) or ADS (50%) manner, which were either semantically congruent (50%) or semantically incongruent (50%) with the priming stimuli. Thus, there were four experimental conditions: ADS-congruent (25%), ADS-incongruent (25%), IDS-congruent (25%), and IDS-incongruent (25%). The second experimenter rubbed participants' right index finger with the priming stimulus within 1000 ms of the presentation of a fixation point on the screen. Following a delay interval ranging between 500 and 600 ms, the auditory target was presented for 650 to 850 ms. After target presentation, two words were presented on the screen; one was the target word which had been presented as auditory stimuli and the other was a distractor word semantically unrelated to the auditory stimuli. Participants were instructed to indicate as quickly and accurately as possible which word on the screen they had just heard presented (Fig. 1). To indicate their decision, participants had to press a button with their left middle or left index finger. The purpose of this task was to ensure that participants actively attended to the target stimuli.

Each experimental session consisted of four blocks, each comprised of 96 trials. Trial order was randomized within each block, with each auditory stimulus presented equally often in combination with a congruent and an incongruent tactile stimulus. In order to ensure that participants understood the procedure, participants performed a practice session composed of six trials before participating in the experimental sessions. The procedure in the practice session was the same as that of the experimental trials except for the absence of priming. All visual stimuli including the fixation point and the prompt for the responses in each trial, as well as the instructions for the task, were presented on a 22-in CRT monitor (RDT223BK, MITSUBISHI). E-prime Software (Psychology Software Tools, Inc., Pittsburgh, PA) was used to present all visual and auditory stimuli and record the participants' responses.

Frequency of Target Word Usage. After the ERP experiment, mothers were again presented with the auditory stimuli (tactile-related words) utilized in the experiment and asked to indicate the frequency with which they used those words in daily life with their children. The following question was answered with numbered scales for each individual auditory stimulus: How often do you (participants) use these words to your baby in everyday life?, 1 (never) to 5 (very often). The total score of the three words presented in the experiment was calculated for each mother.

EEG Data Acquisition and Processing. EEG data were recorded with a 64-channel Geodesic Sensor Net and analyzed using Net Station software (EGI, Eugene, OR) sampled at 250 Hz with a 0.1–100 Hz band-pass filter. Impedances were measured prior to and following EEG recording. Before recording, impedances were below 50 k Ω . All recordings were initially referenced to the vertex and later re-referenced to the average of all channels. In off-line analysis, EEG data were digitally filtered using a 0.3–30 Hz band-pass filter. The data were segmented into 1000 ms epochs time-locked to the onset of the auditory stimulus (target) with a 100 ms pre-stimulus baseline period. Artifacts were screened with automatic detection methods as follows: segments containing eye blink (80 μ V threshold within 20 ms in the frontal region), eye movement artifacts (55 μ V threshold), and channels with amplitudes exceeding ± 80 μ V were excluded from the averaging. Segments including more than ten bad channels were also excluded from averaging. Additionally, EEG records were edited for motor artifacts such as body movement based on visual inspection. The averages of amplitudes were computed separately for each condition (ADS-congruent, ADS-incongruent, IDS-congruent, IDS-incongruent) for each group.

Statistical Analyses. We computed the amplitude across electrodes in the frontal to central regions according to previous research²³ (for the electrode sites analyzed in this study, see Supplementary Figure S1). Mean amplitude for each condition at each time point was calculated. As a preliminary analysis, we conducted ANOVAs with

prosody (2: ADA/IDS) and congruency (2: congruent/incongruent) as within-subjects factors at each time point. To avoid the detection of spurious differences among conditions, we considered a time range of 7 consecutive time points (28 ms) of p -values < 0.05 to indicate a significant effect. We set the following three periods for the analysis: N1 (120–180 ms after stimulus onset), P2 (200–300 ms after stimulus onset), and N400 (300–500 ms after stimulus onset). Mean amplitude in each period was computed for each condition. These variances were analyzed by mixed measures ANOVAs with prosody (2: ADS/IDS), and congruency (2: congruent/incongruent) as within-subjects factors, and group (2: mothers/non-mothers) as a between-subjects factor.

1. Fernald, A. Prosody in speech to children: prelinguistic and linguistic functions. *Ann. Child Dev.* **8**, 43–80 (1991).
2. Trainor, L. J., Clarke, E. D., Huntley, A. & Adams, B. A. The acoustic basis of preferences for infant-directed singing. *Inf. Behav. Dev.* **20**, 383–396 (1997).
3. Soderstrom, M. Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Dev. Rev.* **27**, 501–532 (2007).
4. Barker, B. A. & Newman, R. S. Listen to your mother! The role of talker familiarity in infant streaming. *Cognition* **94**, 45–53 (2004).
5. Cooper, R. P. & Aslin, R. N. Preference for infant-directed speech in the first month after birth. *Child Dev.* **61**, 1584–1594 (1990).
6. Werker, J. F. & McLeod, P. J. Infant preference for both male and female infant-directed talk: a developmental study of attentional and affective responsiveness. *Can. J. Psychol.* **43**, 230–246 (1989).
7. Taumoepeau, M. & Ruffman, T. Stepping stones to others' minds: maternal talk relates to child mental state language and emotion understanding at 15, 24, and 33 months. *Child Dev.* **79**, 284–302 (2008).
8. Kuhl, P. K. & Rivera-Gaxiola, M. Neural substrates of language acquisition. *Annu. Rev. Neurosci.* **31**, 511–534 (2008).
9. Thiessen, E. D., Hill, E. A. & Saffran, J. R. Infant-directed speech facilitates word segmentation. *Infancy* **7**, 53–71 (2005).
10. Vallabha, G. K., McClelland, J. L., Pons, F., Werker, J. F. & Amano, S. Unsupervised learning of vowel categories from infant-directed speech. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 13273–13278 (2007).
11. Ramírez-Esparza, N., García-Sierra, A. & Kuhl, P. K. Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev. Sci.* **17**, 1–12 (2014).
12. Sullivan, J. W. & Horowitz, F. D. Infant intermodal perception and maternal multimodal stimulation: Implications for language development. *Advances in Infancy Research* **2**, 183–239 (1983).
13. Liebal, K., Behne, T., Carpenter, M. & Tomasello, M. Infants use shared experience to interpret pointing gestures. *Dev. Sci.* **12**, 264–271 (2009).
14. Gogate, L. J., Bahrick, L. E. & Watson, J. D. A study of Multimodal Motherese: The Role of Temporal Synchrony between Verbal Labels and Gestures. *Child Dev.* **71**, 878–894 (2000).
15. Gogate, L. J., Bolzani, L. H. & Betancourt, E. A. Attention to maternal multimodal naming by 6- to 8-month-old infants and learning of word-object relations. *Infancy* **9**, 259–288 (2006).
16. Gogate, L. J., Walker-Andrews, A. S. & Bahrick, L. E. The intersensory origins of word-comprehension: An ecological-dynamic systems view. *Dev. Sci.* **4**, 1–18 (2001).
17. Matsuda, Y. et al. Processing of infant-directed speech by adults. *NeuroImage* **54**, 611–621 (2011).
18. Bergelson, E. & Swingle, D. At 6–9 months, human infants know the meanings of many common nouns. *Proc. Natl. Acad. Sci.* **109**, 3253–3258 (2012).
19. Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M. & Stager, C. L. Acquisition of word-object associations by 14-month-old infants. *Dev. Psychol.* **34**, 1289–1309 (1998).
20. Ferber, S. G., Feldman, R. & Makhoul, I. R. The development of maternal touch across the first year of life. *Early Hum. Dev.* **84**, 363–370 (2008).
21. Fernald, A. & Morikawa, H. Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Dev.* **64**, 637–656 (1993).
22. Yoshida, H. A cross-linguistic study of sound symbolism in children's verb learning. *J. Cognition. Dev.* **13**, 232–265 (2012).
23. Schneider, T. R., Debener, S., Oostenveld, R. & Engel, A. K. Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage* **42**, 1244–1254 (2008).
24. Schneider, T. R., Lorenz, S., Senkowski, D. & Engel, A. K. Gamma-band activity as a signature for cross-modal priming of auditory object recognition by active haptic exploration. *J. Neurosci.* **31**, 7, 2502–2510 (2011).
25. Verkindt, C., Bertrand, O., Thevenet, M. & Pernier, J. Two auditory components in the 130–230 ms range disclosed by their stimulus frequency dependence. *NeuroReport* **5**, 1189–1192 (1994).
26. Senkowski, D., Saint-Amour, D., Kelly, S. P. & Foxe, J. J. Multisensory processing of naturalistic objects in motion: A high-density electrical mapping and source estimation study. *NeuroImage* **36**, 877–888 (2007).
27. Liebenthal, E. et al. Specialization along the left superior temporal sulcus for auditory categorization. *Cereb. Cortex* **20**, 2958–2970 (2010).
28. Parkinson, A. L. et al. Understanding the neural mechanisms involved in sensory control of voice production. *NeuroImage* **61**, 314–322 (2012).



29. Parsons, C. E., Stark, E. A., Young, K. S., Stein, A. & Kringelbach, M. L. Understanding the human parental brain: A critical role of the orbitofrontal cortex. *Soc. Neurosci* **8**, 525–543 (2013).
30. Kringelbach, M. L. & Rolls, E. T. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Prog. Neurobiol.* **72**, 341–372 (2004).
31. Tremblay, K. N. & Kraus, N. Auditory training induces asymmetrical changes in cortical neural activity. *J Speech Lang Hear Res.* **45**, 564–572 (2002).
32. Kutas, M. & Federmeier, K. D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Ann. Rev. Psychol.* **62**, 621–647 (2011).
33. Orgs, G., Lange, K., Dombrowski, J. H. & Heil, M. N400-effects to task-irrelevant environmental sounds: Further evidence for obligatory conceptual processing. *Neurosci. Lett.* **436**, 133–137 (2008).
34. Kiefer, M. Repetition-priming modulates category-related effects on event-related potentials: further evidence for multiple cortical semantic systems. *J. Cognitive Neurosci.* **17**, 199–211 (2005).

Acknowledgments

This study was supported by funding from the Japan Science and Technology Agency, Exploratory Research for Advanced Technology, Okanoya Emotional Information Project, and Grants-in-Aid for Scientific Research from the Japan Society for the Promotion of Science and the Ministry of Education Culture, Sports, Science and Technology (24119005, 24300103 to M.M.-Y.; 13J05878 to Y.T.). We would like to thank all the children and parents who participated in the research. We also thank N. Naoi, and Y. Fuchino for

technical support, and H. Hirai, S. Mizugaki, M. Yamamoto, and Y. Nishimura, for their assistance in the experiment.

Author contributions

Y.T., K.O. and M.M.-Y. designed the research. Y.T. and H.F. conducted experiments and analyzed the data. Y.T., H.F. and M.M.-Y. wrote the main manuscript text. Y.T., H.F., K.O. and M.M.-Y. reviewed and discussed the main manuscript, and approved the final manuscript.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tanaka, Y., Fukushima, H., Okanoya, K. & Myowa-Yamakoshi, M. Mothers' multimodal information processing is modulated by multimodal interactions with their infants. *Sci. Rep.* **4**, 6623; DOI:10.1038/srep06623 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>