By the sound of it. An ERP investigation of human action sound processing in 7-month-old infants

Elena Geangu [a,*], Ermanno Quadrelli [b], James W. Lewis [c],
Viola Macchi Cassia [b,1], Chiara Turati [b]

[a] Department of Psychology, Lancaster University, Bailrigg, Lancaster LA1 4YF, United Kingdom
[b] Department of Psychology, University of Milano-Bicocca, Milano, 20126, Italy
[c] Department of Neurobiology & Anatomy, and Center for Advanced Imaging, West Virginia University, Morgantown, WV 26506, USA

ARTICLE INFO

ABSTRACT

Recent evidence suggests that human adults perceive human action sounds as a distinct category from human vocalizations, environmental, and mechanical sounds, activating different neural networks (Engel et al., 2009; Lewis et al., 2011). Yet, little is known about the development of such specialization. Using event-related potentials (ERP), this study investigated neural correlates of 7-month-olds' processing of human action (HA) sounds in comparison to human vocalizations (HV), environmental (ENV), and mechanical (MEC) sounds. Relative to the other categories, HA sounds led to increased positive amplitudes between 470 and 570 ms post-stimulus onset at left anterior temporal locations, while HV led to increased negative amplitudes at the more posterior temporal locations in both hemispheres. Collectively, human produced sounds (HA + HV) led to significantly different response profiles compared to non-living sound sources (ENV + MEC) at parietal and frontal locations in both hemispheres. Overall, by 7 months of age human action sounds are being differentially processed in the brain, consistent with a dichotomy for processing living versus non-living things. This provides novel evidence regarding the typical categorical processing of socially relevant sounds.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

For most humans hearing is essential to everyday social life. It enables communication through language and provides information about the dynamics of the social agents in our close environment. Perceiving fine changes in the pitch of someone's voice can tell us whether they are disappointed or happy (Grossmann et al., 2005), while the rhythm and intensity of footstep sounds provide us with information about a social agent's speed of motion and

emotional state (Sievers et al., 2013). Sound onset and offset also coincide with the occurrence of visual cues for motion, supporting the efficient segmentation of others' actions in meaningful units (Thomas and Shiffrar, 2010). Although extensive evidence has shown that vocalizations are perceived by humans as a distinct category of natural sounds (Belin et al., 2004; Fecteau et al., 2004), and that specialization for vocal processing emerges very early in infancy (Blasi et al., 2011), less is known about perceiving sounds associated with human actions (Lewis et al., 2005, 2006, 2011). Importantly, we have little to no understanding of how this ability develops in the first years of life. The main aim of the current study is to increase our knowledge in this respect by investigating the neural correlates of infants' perception of human action sounds.

* Corresponding author. Tel.: +44 1524 593565.
E-mail address: e.geangu@lancaster.ac.uk (E. Geangu).
[1] For the fourth author, Macchi Cassia is the double family name.

At birth, infants already have incipient abilities for auditory localization (Morrongiello and Clifton, 1984) and for segregating concurrent sound streams leading to the identification of objects in the environment (Winkler et al., 2003). The origins of these abilities can be traced back to prenatal life, since fetuses detect low frequency sounds (250–500 Hz) by 19 weeks gestational age and show a gradual increase in the ability to detect high frequency sounds (1000–3000 Hz) after 35 weeks (Burnham and Mattock, 2010). The early maturation of the auditory system and in-utero access to social sounds (i.e., mother's voice, native language) are probably linked to new-borns' ability to discriminate speech and non-speech vocalizations. For instance, new-borns strongly prefer their mother's voice compared to a stranger's voice (Moon and Fifer, 1986) and the sound rhythm specific to their native language compared to that of other natural languages (Kuhl, 2010; Nazzi et al., 1998).

More recently, neurophysiological studies have shown that processing human voice sounds specifically activates the left hemisphere of the new-born brain (Pena et al., 2003) and distinct cerebral networks in few-month-old infants (Grossmann et al., 2010). Such specialization is not confined to speech (Benavides-Varela et al., 2011; Dehaene-Lambertz et al., 2002; Peña et al., 2010), but also extends to non-speech human vocalizations (Blasi et al., 2011; Lloyd-Fox et al., 2011). Using functional near-infrared spectroscopy (fNIRS), Lloyd-Fox et al. (2011) showed that non-speech vocalizations (i.e., crying, laughing, yawning, coughing) induce greater activation of the bilateral anterior temporal cortex compared to non-human sounds (i.e., toys rattling, water running), a difference which becomes more prominent between 4 and 7 months of age. Similar results were found in response to non-speech voice in comparison to environmental sounds, with 3–7-month-olds showing greater activation for vocalizations in the right anterior middle and superior temporal gyri (Blasi et al., 2011), which overlaps with the voice sensitive regions in the adult brain (Belin et al., 2000).

Although there is a general consensus that voices play a prominent role in infants' early social and cognitive development, less attention has been devoted to the investigation of whether and how other types of human-produced sounds may affect infants' understanding of their social environment. There is some evidence that social input provided by voices affects infants' processing of other social information, including that related to others' actions (Brand and Tapscott, 2007). Nonetheless, in the home environment adults move in infants' proximity doing their daily chores, and while some of this motion is associated with vocalizations, it is almost always accompanied by the sounds generated by human actions. For adults, these sounds are as important for processing others' actions as visual information is (Alaerts et al., 2011). Action sounds can be sufficient for adults to discriminate between actions performed by another individual (Keysers et al., 2003) and influence how we perceive visual information about actions. When participants are asked to judge the gender of a point-light display walker, the sound of female footsteps increases the likelihood that this will be perceived as being female (Van der Zwan et al., 2009). Also, the

presence of action sounds like footsteps, but not other synchronous sounds, increases visual sensitivity in detecting point-light displays of coherent human gait (Thomas and Shiffrar, 2010).

Like vocalizations, action sounds selectively activate specific cortical networks in the adult brain. A series of fMRI studies (Engel et al., 2009; Lewis et al., 2011) have shown that human action sounds preferentially activate those brain areas which have been functionally linked to processing others' actions as goal directed (i.e., left inferior frontal gyrus—IFG, inferior parietal lobule—IPL, and primary motor cortex—M1), as well as those areas typically associated with processing of biological motion cues from different sensory modalities (i.e., bilateral posterior superior temporal sulcus—pSTS and posterior middle temporal gyrus—pMTG). In contrast, sounds produced by mechanical devices elicit selective activation of brain areas associated with high-level stages of acoustic signal processing (i.e., the anterior superior temporal gyrus—aSTG and the left parahippocampal cortex) as well as with the processing of visual information conveyed by objects and scenes (Engel et al., 2009; Lewis et al., 2011). Other sounds encountered in our natural environment activate the areas involved in integrating congruent intermodal features of audio-visual stimuli. At the level of the scalp, human action sounds trigger positive ERPs at left temporal and central locations associated with the activity of STS, while non-biological sounds elicit bilateral positive ERP activity at temporo-parietal and frontal locations associated with bilateral temporal activation (Pizzamiglio et al., 2005).

Despite the relevance that human action sounds have for adults' social understanding, no study up to now has investigated infants' ability to process human action sounds in the earliest stages of postnatal life. The aim of the current study was to investigate this issue by examining the neural correlates of infants' processing of sounds produced by human actions in comparison to human non-speech vocalizations and non-biological sounds.

Electroencephalography (EEG) and ERPs have high temporal accuracy and have been successfully used to investigate infants' processing of visual and auditory social information (De Haan, 2007; Trainor, 2010). For example, ERPs show that 4-month-old infants discriminate between their mother's voice and the voice of a stranger (Purhonen et al., 2005), and are sensitive to the development of infants' ability to detect changes in auditory characteristics, like the pitch, important for processing socially relevant sounds (Kushnerenko et al., 2002). These findings render ERPs a promising method for investigating whether human action sounds are processed distinctly from non-biological sounds and whether human action sounds and human non-speech vocalizations represent distinct sub-categories of human produced sounds.

We chose to test 7-month-old infants as previous studies have shown that by this age auditory cerebral specialization for speech and non-speech human vocalizations is already present (Blasi et al., 2011; Grossmann et al., 2010; Lloyd-Fox et al., 2011). Infants were presented in a random order with four sound categories: sounds produced by human actions (Human Action, HA), human non-speech vocalizations with neutral emotional prosody

(Human Vocalization, HV), sounds produced by automated mechanical devices (Mechanical, MEC), and sounds produced by nature (Environmental, ENV). We investigated whether and how infant ERP responses differ between HA and other types of sounds (HV, MEC, and ENV), between human produced sounds (HA and HV) and non-human produced sounds (MEC and ENV); and between HV and other types of sounds (HA, MEC, ENV). Based on previous studies using similar stimuli and paradigm with infants and adults, we hypothesized that infants' enhanced processing of HA would be reflected in a larger amplitude of the ERP late slow wave (LSW) component at temporal sites (Grossmann et al., 2005; Grossmann and Johnson, 2007). In order to determine whether HA and HV are processed as instances of the global category of human produced sounds, we analyzed ERP responses recorded at frontal and parietal locations. A larger LSW for both HA and HV compared to MEC and ENV sounds was expected to be recorded at frontal electrodes (Guillem et al., 2001; Quinn et al., 2006), while an increased negativity in ERP responses to human (HA, HV) compared to non-human action sounds (MEC, ENV) was expected at parietal locations (Hirai and Hiraki, 2005; Marshall and Shipley, 2009). Finally, although no previous study has investigated infant ERP responses to neutral non-speech human vocalizations compared to non-vocalization sounds, extant evidence on infant ERP correlates of voice processing suggests specific activation at fronto-central and temporal locations within both hemispheres (Cheng et al., 2012; Grossmann et al., 2005).

2. Methods

2.1. Participants

Fifteen 7-month-old infants (6 females, $M = 214$ days, SD = 8 days) were included in the final analyses. All infants were born full-term (37–42 weeks gestation), had a normal birth weight (>2500 g), did not suffer of any neurological or other medical conditions, and were observed to have normal vision for their age. Eight additional infants were tested but not included in the final sample due to fussiness ($n = 3$), excessive artifacts ($n = 4$), or technical problems with data collection ($n = 1$). The procedure followed the ethical standards (the Declaration of Helsinki, BMJ 1991; 302:1194) and was approved by the University ethical committee. Parents gave written informed consent for their infants' participation.

2.2. Stimuli

Participants were presented with audio recordings of HA (i.e., footsteps, clapping, zippering), HV (i.e., yawning, sighing, coughing) depicting relatively non-communicative and emotionally neutral vocal expressions, ENV (i.e., wind, rain, water flowing), and MEC (i.e., engine start, micro-wave oven, washing machine) sounds. Each category included 9 different exemplars. The auditory stimuli were extracted from or modified based upon an existing database of category specific sounds (Engel et al., 2009). They were screened by 6 adult listeners and edited to 1000 ms duration. All of the sound stimuli were

assessed for various "bottom-up" spectro-temporal signal attributes (Table 1) that may correlate with their "top-down" conceptual boundaries or anticipated ERP response differences. Sounds were initially balanced psychophysically for perceived overall loudness. The intensity of the sounds was then assessed quantitatively for total Root Mean Square (RMS) power using Adobe Audition 3 (Adobe Systems Inc.) and balanced to roughly −28 dB RMS power across all four categories. A pitch analysis was performed for each sound based on an autocorrelation method using Praat software (http://www.fon.hum.uva.nl/praat/), with standard default settings (75 Hz pitch floor and 600 Hz ceiling). Some sounds did not have a quantifiable estimated pitch (Table 1, not available pitch entries—NA). Using methods described earlier (Lewis et al., 2005, 2009, 2012), we derived a Harmonics-to-Noise Ratio (HNR) mean value, a measure of spectro-temporal dynamics based on entropy (Wiener entropy) in the acoustic signal (Tchernichovski et al., 2001), plus a measure of spectral structure variance (SSV) of the sounds. The SSV is a measure of changes in signal entropy over time that has been shown to have utility in categorizing natural sound signals (Reddy et al., 2009).

2.3. Procedure

ERPs were recorded while infants sat on their parent's lap at a distance of approximately 60 cm from a 24-inch monitor in a dimly lit, audiometric and electrically shielded cabin. The auditory stimuli were randomly presented using E-Prime software v2.0 (Psychology Software Tools Inc., Pittsburgh, PA) over two speakers placed to the left and right of the monitor. No more than two sounds from the same category were presented consecutively. A trial consisted of 1000 ms stimulus presentation followed by a silent interstimulus which varied randomly in duration between 900 and 1100 ms. A black fixation cross on a grey background was displayed on the monitor during the entire stimulus presentation. The procedure continued until maximum 288 trials were presented (72/category) or until the infant became bored. In order to minimize infants' movements and maintain their interest, an experimenter was present to quietly blow soap bubbles throughout the entire procedure. Both the parent and the experimenter avoided visual contact and any attempt for social interaction with the participant. Stimulus presentation lasted approximately 12 min.

2.4. Electroencephalogram recording and data analysis

The electroencephalogram (EEG) was recorded continuously using a 128-electrode HydroCel Geodesic Sensor Net (Electrical Geodesic Inc., Eugene, OR) and amplified using an EGI NetAmps 300 amplifier. The signal was referenced online to the vertex electrode (Cz), a bandpass filter of .1–100 Hz was applied, and the data were sampled at 500 Hz. Impedances were checked prior to the beginning of recording and considered acceptable if lower than 50 kΩ. EEG data were further processed offline using NetStation v4.6.4 (Eugene, OR). The signal was band-pass filtered (0.3–30 Hz), and the ERP trials were segmented with a 100 ms baseline and 1000 ms following stimulus onset.

Sound category	Sound stimuli		Amplitude	Mean Pitch	Harmonics-to-noise ratio (HNR) Mean	Wiener entropy (WE)	Spectral structure variation (SSV)
			Total RMS power (dB)	Pitch (Hz)	HNR (dB)	WE	SSV
Human Actions (HA)	Footsteps	HA Footsteps 1	−35.78	NA	−2.21	−1.17	1.39
		HA Footsteps 2	−27.12	NA	0.16	−3.98	1.17
		HA Footsteps 3	−27.30	NA	3.92	−3.10	2.10
	Hand clapping (one individual)	HA Handclapping 1	−28.47	NA	1.14	−3.95	1.22
		HA Handclapping 2	−24.88	NA	−1.99	−6.86	1.58
		HA Handclapping 3	−30.11	NA	1.39	−4.10	3.18
	Zippering	HA Zippering 1	−28.28	263	−0.16	−4.07	0.31
		HA Zippering 2	−28.38	187	−0.30	−4.07	0.31
		HA Zippering 3	−28.41	375	−0.61	−1.97	0.34
		HA Average =	**−28.75**	**275**	**0.15**	**−3.70**	**1.29**
Non-verbal Human Vocalizations (HV)	Cough	HV Female Cough 1	−28.11	331	2.76	−3.44	0.89
		HV Female Cough 2	−29.35	476	3.02	−6.23	2.05
		HV Male Cough 1	−29.87	323	3.44	−5.97	1.10
	Yawn	HV Male Yawn 1	−29.04	182	15.10	−6.71	2.29
		HV Male Yawn 2	−28.26	248	2.79	−5.69	1.61
		HV Male Yawn 3	−28.11	191	6.92	−6.38	3.05
	Sigh	HV Female Sigh 1	−28.58	337	2.22	−5.29	0.47
		HV Female Sigh 2	−28.55	322	11.08	−4.92	1.29
		HV Female Sigh 3	−28.06	322	7.05	−6.47	1.49
		HV Average=	**−28.66**	**304**	**6.04**	**−5.68**	**1.58**
Mechanical (ME)	Engine	ME Engine 1	−28.31	503	2.30	−4.10	1.00
		ME Engine 2	−19.87	549	6.40	−7.83	1.15
		ME Engine 3	−28.38	NA	−2.05	−7.92	0.26
		ME Engine 4	−28.10	NA	0.03	−3.49	0.79
	Microwave	ME Microwave 1	−28.02	175	2.07	−11.29	0.58
		ME Microwave 2	−28.00	89	4.10	−7.04	0.48
	Washing machine & Vacuum Cleaner	ME Vacuum Cleaner 1	−28.00	213	15.58	−6.28	0.36
		ME Washing Machine 1	−28.27	99	1.69	−6.81	0.97
		ME Washing Machine 2	−28.41	156	0.82	−4.99	0.46
		ME Average=	**−27.26**	**255**	**3.44**	**−6.64**	**0.67**
Environment (EN)	Rain&Water bubbling	EN Rain 1	−30.32	NA	−3.30	−1.89	0.18
		EN Rain 2	−30.54	111	1.76	−4.30	0.64
		EN Water Bubbling 1	−27.80	427	6.10	−8.39	8.69
	Wind	EN Wind 1	−24.13	206	1.30	−7.68	0.60
		EN Wind 2	−24.96	231	1.07	−11.65	0.66
		EN Wind 3	−28.25	268	0.44	−11.31	0.46
	River	EN River 1	−30.28	NA	−3.84	−2.64	0.23
		EN River 2	−28.49	406	−1.43	−3.65	0.60
		EN River 4	−28.34	NA	−3.04	−5.08	0.27
		EN Average=	**−28.12**	**275**	**−0.11**	**−6.29**	**1.37**

Note. All sound stimuli had duration = 1000.0 ms for all stimuli, intensity matched at roughly −28 dB total RMS power, on/off ramped 25 m. NA = not available pitch.
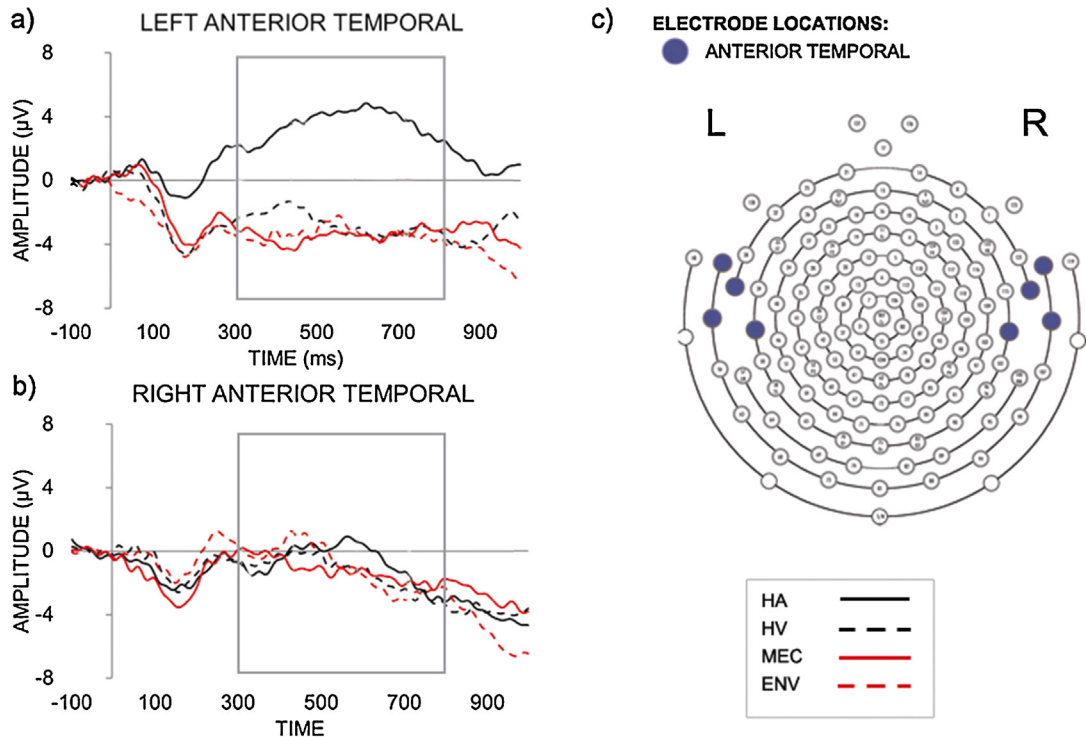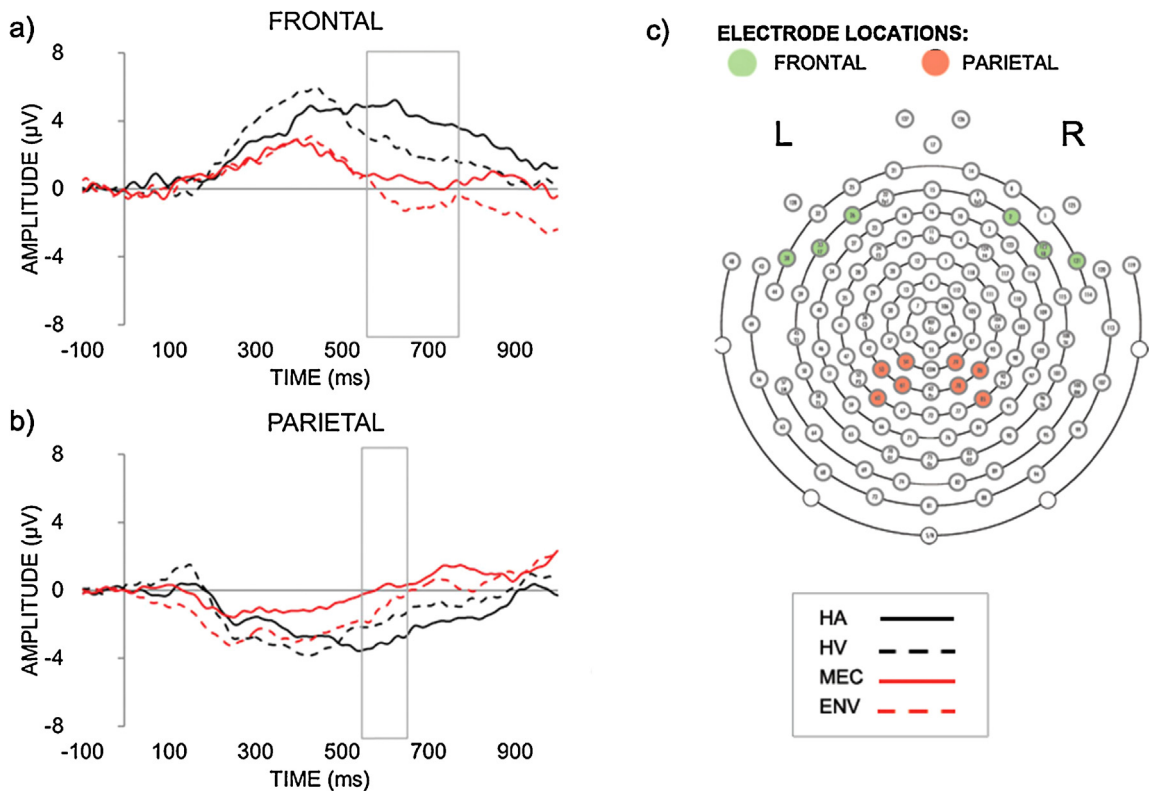
Fig. 1. Waveform plots depicting grand-average ERPs in response to HA (solid black line), HV (dash black line), MEC (solid red line), ENV (dash red line) over the right (A) and left (B) anterior temporal areas. (C) Electrode groupings for the anterior temporal areas consisting of channels 108 (T4), 113, 114, 120 on the right and 45 (T3), 43, 44, 49 on the left. Please note that the positive is plotted upwards. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Data were corrected to the average voltage during baseline and re-referenced to the algebraic mean of all channels. To eliminate artifacts, segmented data were automatically rejected whenever the signal exceeded $\pm 200\,\mu V$ at any electrode. Data were further checked through visual inspection for eye-movements, eye-blinks and other body movement artifacts not detected by the automated algorithm. Trials were excluded if more than eighteen bad channels were detected. Of the remaining trials, individual bad channels were replaced using spherical spline interpolation. Individual subject averages were computed separately for each channel across all trials within each condition and then re-referenced to the average reference. Across participants, the mean number of trials contributing to the average ERP was 34.35 (HA: 35, SD = 11.5; HV: 34.6, SD = 10.8; ENV: 34, SD = 11.1; MEC: 33.8, SD = 10.2). A similar number of trials contributed to the final analysis for each condition, $F(3,42) = .203$ $p > .80$. Because this was the first study to investigate infant ERPs in response to human action sounds and emotionally neutral non-verbal vocalization using random stimulus presentation, we were unable to formulate predictions regarding the specific timing and polarity of all ERP components. Accordingly, like in other infant ERP studies (e.g., Quinn et al., 2006), our analyses were based on visual inspection of the waveforms, which revealed several time windows and scalp regions of interest (ROI). In line with previous studies (Grossmann et al., 2005), a difference in the amplitude of the late slow wave (LSW) was observed between 300 and 800 ms from

stimulus onset at anterior temporal locations, giving rise to an anterior positive slow wave—PSW (Fig. 1), and a difference in the amplitude of the NSW was observed at posterior temporal ROI between 470 and 570 ms (Fig. 3). A positive component (Pc) was observed in the frontal ROI, with maximum amplitude between 300 and 500 ms as previously described (Friederici et al., 2007; Parise et al., 2010). A PSW with maximum differentiation between conditions within 550 and 750 ms from stimulus onset was evident at the frontal ROI (Fig. 2) and within 550–730 ms at the central ROI. Similar to previous studies on biological motion (Hirai and Hiraki, 2005; Marshall and Shipley, 2009), a NSW was observed at parietal ROI (550–650 ms), which differentiated between human and non-human produced sounds (Fig. 2). All ERP components included in the analysis were analyzed as average amplitudes, averaging all datapoints within the given time window. The same procedure was also used for the Pc component. Due to the fact that at least one condition did not show a clear peak voltage, the average amplitude for the entire window rather than the peak voltage was included in the analysis.

ERP mean amplitudes within each region of interest were analyzed using 4 (category: HA, HV, ENV, MEC) by 2 (hemisphere: left and right) repeated measures ANOVAs. All statistical tests were conducted at .05 level of significance (two-tailed), and paired-sample t-tests were corrected for multiple comparisons using the Holm-Bonferroni stepwise procedure.

Fig. 2. Waveform plots depicting grand-average ERPs in response to HA (solid black line), HV (dash black line), MEC (solid red line), ENV (dash red line) over the frontal (A) and parietal (B) areas. (C) Electrode groupings for the frontal area (green) consisting of channels 26, 33 (F7), 38, 2, 121, 122 (F8) and for the parietal area (red) consisting of channels 53, 54, 60, 61, 78, 79, 85, 86. Please note that the positive is plotted upwards. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3. Results

3.1. Human action (HA) vs. other types of sounds (HV, MEC, and ENV)

3.1.1. Anterior temporal electrodes

Anterior temporal electrodes recorded a PSW occurring within a 300–800 ms window, suggesting a differentiation between the HA sounds and the other types of sounds.

A 4×2 repeated measures ANOVA on the mean amplitude of this PSW showed a significant effect of condition, $F(3,42) = 4.650$, $p = .022$, $\eta^2 = .248$, which was qualified by a significant Condition × Hemisphere interaction, $F(3,42) = 4.979$, $p = .001$, $\eta^2 = .248$. Post-hoc pairwise comparisons showed that at left temporal locations HA sounds elicited increased mean positive amplitude ($M = 3.563 \,\mu V$, $SD = 3.735$) compared to HV ($M = -2.626 \,\mu V$, $SD = 5.766$), $t(14) = 5.805$, $p < .001$, MEC ($M = -3.459 \,\mu V$, $SD = 3.998$), $t(14) = 5.573$, $p < .001$, and ENV ($M = -3.370 \,\mu V$, $SD = 3.749$), $t(14) = 5.805$, $p < .001$, sounds (Fig. 1). HA also elicited a higher amplitude of the PSW in the left ($M = 3.564 \,\mu V$, $SD = 3.735$) compared to the right ($M = -.689 \,\mu V$, $SD = 3.249$) hemisphere, $t(14) = 3.336$, $p = .035$. All other differences were not significant ($p > .258$).

3.1.2. Human produced (HA, HV) vs. non-human produced (MEC, ENV) sounds

The average amplitudes recorded at both parietal and frontal electrodes within different time windows suggested a differentiation between human (HA, HV) and non-human (MEC, ENV) produced sounds.

3.1.3. Parietal electrodes

A 4×2 repeated measures ANOVA on the mean amplitude of the NSW recorded within the 550–650 ms window in the parietal ROI showed a significant effect of condition, $F(3,42) = 3.402$, $p = .026$, $\eta^2 = .196$. Pairwise comparisons suggested that HA sounds elicited bilaterally a larger negativity ($M = -3.068 \,\mu V$, $SD = .971$) compared to MEC ($M = .158 \,\mu V$, $SD = 1.091$, $p = .032$) and ENV ($M = -.970 \,\mu V$, $SD = .746$, $p = .035$) sounds but not to HV sounds ($M = -1.795 \,\mu V$, $SD = 1.099$, $p = .232$) (Fig. 2). All other comparisons were non-significant ($p > .09$).

3.1.4. Frontal electrodes

The 4×2 repeated measures ANOVA on the average amplitude of the frontal Pc (300–500 ms) revealed a main effect of condition, $F(1,42) = 4.059$, $p = .013$, $\eta^2 = .225$. The pairwise comparisons revealed that HV elicit a significantly higher amplitude ($M = 5.271 \,\mu V$, $SD = .873$) compared to MEC sounds ($M = 2.433 \,\mu V$, $SD = .782$), $p = .025$. However,
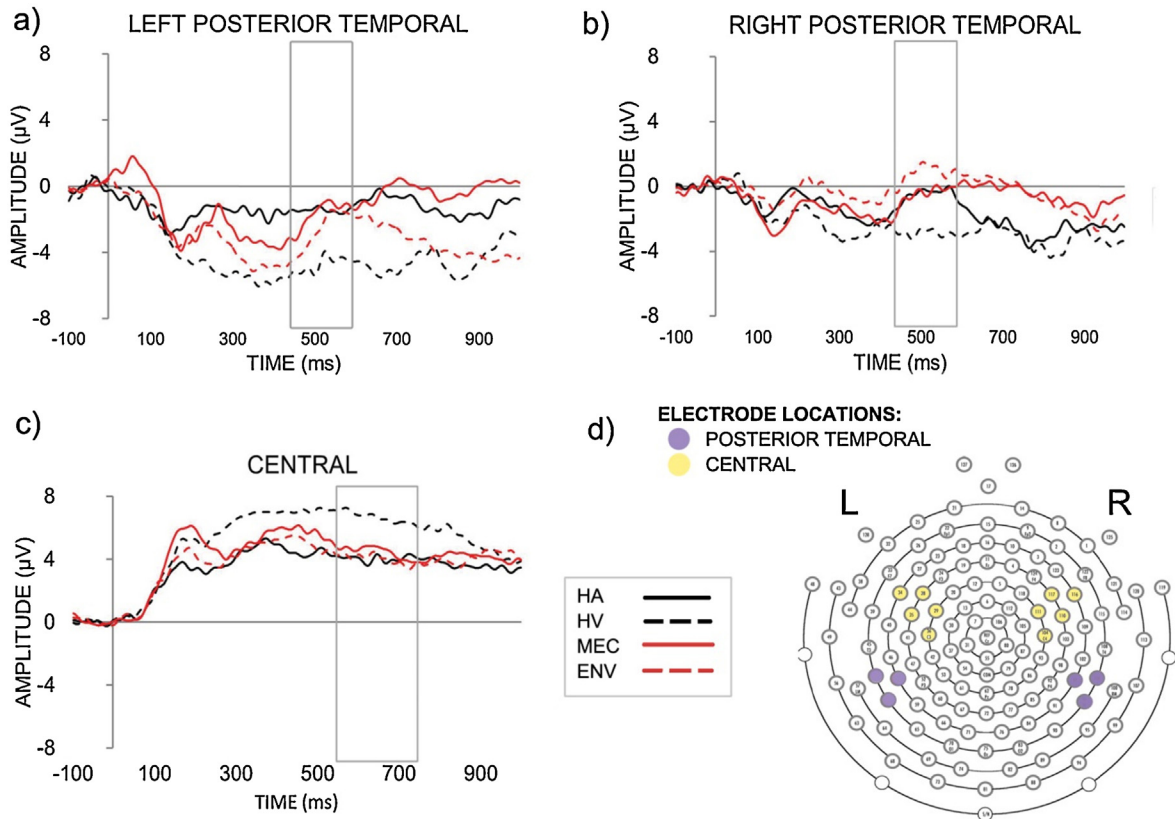
Fig. 3. Waveform plots depicting grand-average ERPs in response to HA (solid black line), HV (dash black line), MEC (solid red line), ENV (dash red line) over the right (A) and left (B) posterior temporal areas and over the central area (C). (D) Electrode groupings for the central area (yellow) consisting of channels 28, 29, 34, 35, 36 (C3), 111, 117, 104 (C4), 110, 116 and for the posterior temporal area (violet) consisting of channels 96 (T6), 97, 101 on the right and 50, 51, 58 (T5) on the left. Please note that the positive is plotted upwards. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the main effect of condition was also qualified by a significant Condition × Hemisphere interaction, $F(1,42) = 3.396$, $p = .026$, $\eta^2 = .195$. In order to disentangle this interaction, post-hoc paired t-tests were performed, with the significance threshold corrected for multiple comparisons. The results of the analysis revealed that HA sounds elicit a larger Pc amplitude on the left ($M = 6.126 \, \mu V$, $SD = 4.962$) compared to the right ($M = 1.93 \, \mu V$, $SD = 2.832$) hemisphere, $t(14) = 3.517$, $p = .048$. Although HA and HV tended to show higher average amplitude of the Pc on the left hemisphere compared to the ENV and MEC sounds, these differences did not reach the corrected significance threshold ($p > .09$). No other post-hoc comparisons were significant. A 4×2 repeated measures ANOVA on the average amplitude of the frontal PSW (550–750 ms) showed a significant effect of condition, $F(3,42) = 5.788$, $p = .001$, $\eta^2 = .327$. Pairwise comparisons revealed that HA sounds elicited bilaterally a larger amplitude ($M = 4.475 \, \mu V$, $SD = .930$) compared to MEC ($M = .501 \, \mu V$, $SD = 1.167$, $p = .005$) and ENV ($M = -.685 \, \mu V$, $SD = 1.042$, $p = .001$) sounds but not to HV ($M = 4.475 \, \mu V$, $SD = .930$, $p = .071$) sounds. HV also seemed to trigger a larger average amplitude ($M = 2.310 \, \mu V$, $SD = 1.172$) compared to MEC sounds ($M = -.970 \, \mu V$, $SD = .746$, $p = .055$), but

this fell short of significance. All other comparisons were non-significant ($p > .118$) (Fig. 2).

3.2. Human vocalizations (HV) vs. other types of sounds (HA, MEC, ENV)

A differentiation between human vocalization (HV) and the other types of sounds was suggested by the average amplitudes recorded at parietal and central electrodes within different time windows.

3.2.1. Posterior temporal electrodes

A 4×2 repeated measures ANOVA on the NSW mean amplitude (470–570 ms) recorded at posterior temporal locations showed significant main effects of condition, $F(3,42) = 5.221$, $p = .004$, $\eta^2 = .272$, and hemisphere, $F(3,42) = 6.834$, $p = .020$, $\eta^2 = .328$. Post-hoc pairwise comparisons showed that the negative component was larger in the left ($M = -2.461$, SD = 1.011) compared to the right ($M = -.610 \, \mu V$, $SD = .551$) hemisphere ($p = .020$). Moreover, on average, the amplitude of the component across the two hemispheres was larger for HV ($M = -3.764 \, \mu V$, $SD = 1.014$) compared to HA ($M = -.878 \, \mu V$, $SD = .952$, $p = .016$), MEC ($M = -.837 \, \mu V$, $SD = .814$, $p = .001$), and ENV ($M = -.663 \, \mu V$,

$SD = .908$, $p = .004$) sounds (Fig. 3). All other comparisons were non-significant ($p > .828$).

3.3. Central electrodes

The ANOVA on the average amplitude of the PSW (550–730 ms) recorded at central locations revealed a main effect of condition, $F(3, 42) = 3.428$, $p = .026$, $\eta^2 = .197$. On average, infants showed a bilateral increased positivity for HV ($M = 6.705\,\mu V$, $SD = .804$) compared to HA ($M = 4.016\,\mu V$, $SD = .783$, $p = .006$), MEC ($M = 4.523\,\mu V$, $SD = .682$, $p = .029$), and ENV ($M = 4.043\,\mu V$, $SD = 1.297$, $p = .026$) sounds (Fig. 3). All other comparisons were non-significant ($p > .490$).

4. Discussion

The present study compared 7-month-old infants' ERP responses to living versus non-living categories of sound, including human action sounds, human non-speech vocalizations, sounds produced by the environment, and sounds produced by automated mechanical devices. The aim was to explore whether the sounds of human actions are processed distinctly from other social and non-social sounds that infants typically encounter in their natural environment, and whether action sounds and non-speech vocalizations represent distinct sub-categories of human produced sounds. Results suggest that, by the age of 7-months, infants manifest sensitivity to human action sounds, as shown by the specific ERP modulation identified at anterior left temporal locations. Also, a specific ERP modulation was present for human vocalizations at posterior temporal electrodes across both hemispheres, which was absent for human action sounds. Thus, at a certain level, human action sounds and human vocalizations are differentially processed in the brain. Our results further indicate that the two subcategories of human sounds are possibly treated as exemplars of a broader category of human produced sounds ("living things"), as shown by the undifferentiated modulation they elicited at frontal and parietal scalp locations.

Human action sounds triggered larger PSW amplitudes at anterior left temporal locations compared to the other types of sounds. The LSW amplitude is considered to reflect enhanced sensory processing (de Haan and Nelson, 1997; Grossmann et al., 2005), suggesting that the sounds associated with human actions undergo more extensive sensorial processing in comparison to other sounds in 7-month-old infants. ERPs recorded at temporal locations have also been found to be sensitive to adult processing of human action sounds and to be linked to the STS and the premotor cortex activation (Pizzamiglio et al., 2005). These brain areas are part of the network specialized for processing auditory information related to human actions (Engel et al., 2009; Lewis et al., 2011). Although our study does not provide evidence for the neural network driving the anterior temporal PSW, our results speak in favor of a specific perceptual processing of human action sounds in 7-month-old infants. In fact, it is important to note that these anterior temporal effects did not extend to human vocalizations, which, instead, induced a selective modulation of ERP responses at posterior temporal scalp locations.

Several previous studies have shown that by the age of 7-months, the voice region of the temporal cortex manifests selective activation to human speech and non-speech vocal sounds (Blasi et al., 2011; Grossmann et al., 2010; Lloyd-Fox et al., 2011). Our results provide converging evidence, by showing larger NSW amplitude across both hemispheres in response to emotionally neutral non-speech vocalizations compared to non-voice sounds. This increased NSW amplitude suggests that 7-month-old infants manifest increased sensory processing of emotionally neutral non-speech vocalizations compared to non-voice sounds, which could potentially rely on a voice specific brain network. The human voice features extracted this way may be further integrated into category specific representations, as suggested by the PSW effects identified at central locations (Guillem et al., 2001; Kusak et al., 2000; Wilding et al., 1995).

In contrast to previous ERP studies of infant processing of emotional human speech (Grossmann et al., 2005), we report ERPs with negative polarity at a more posterior temporal location. Also, the average amplitude of the Pc signal recorded at frontal electrodes suggests that, unlike for other types of speech sounds (i.e., their mother's voice (Purhonen et al., 2005), speech with negative emotional prosody (Grossmann et al., 2005)), 7-month-old infants in our study did not manifest increased processing or attention to non-speech vocalizations. These differences could be due to the nature of stimuli: speech vs. non-speech, neutral vs. emotional prosody. Imaging methods providing higher spatial resolution (i.e., fMRI—Blasi et al., 2011; fNIRS—Grossmann et al., 2010) have shown that emotional voice stimuli modulate the activation of several brain areas in addition to the temporal cortex (e.g., the insula, the orbitofrontal cortex, the inferior frontal cortex). While we cannot provide a conclusive interpretation of the cortical sources of the effects observed in our study, it is possible that the differences in polarity and scalp location between non-speech vocalizations and speech associated with emotional prosody are due to differences in the brain areas involved in processing these types of stimuli. Although human actions and human non-speech vocalizations in the current study elicited distinct patterns of temporal ERP responses, there was also a certain degree of overlap, particularly for the frontal and the parietal ERP components, which is probably due to the fact that the two sound categories share important features related to the biological source of the sounds.

Compared to the non-living sounds (ENV and MEC), human action sounds elicited larger NSW amplitude bilaterally at parietal scalp locations, suggesting that infants process the auditory information associated with human actions by using mechanisms more generally involved in the processing of information about biological agents. This is generally consistent with human lesion and neuropsychological studies indicating that semantic knowledge relating to different categories of objects, including the distinction of "living" vs "non-living things" (Hillis and Caramazza, 1991; Silveri et al., 1997; Warrington and Shallice, 1984), is represented along partially segregated systems in the brain (Grossman et al., 2002; Glushko et al., 2008; Moore and Price, 1999). The parietal NSW has been

shown to manifest sensitivity to visual information about biological human motion in 5- and 8-month-old infants, and adults (Hirai et al., 2003; Hirai and Hiraki, 2005; Marshall and Shipley, 2009). Similar to the infant ERP responses to point light displays depicting different types of canonical human motion (Marshall and Shipley, 2009), in our study, the auditory characteristics of different human actions elicited an extended NSW, distributed across small clusters of parietal electrodes in both hemispheres. This pattern of responses might reflect the extensive processing required by the presentation of multiple exemplars of human action sounds (Hirai and Hiraki, 2005; Marshall and Shipley, 2009).

We have also recorded greater amplitude of the frontal PSW in response to human action compared to non-living sounds, an effect which was present in both hemispheres. Frontal PSW has been previously documented as a neural correlate of global-level category formation in 6-month-old infants (Quinn et al., 2006), and also associated in adults with the integration of stimuli features into distinctive representations, updating these representations with newly acquired information, and the retrieval of this information (Guillem et al., 2001; Kusak et al., 2000; Wilding et al., 1995). It is thus possible that the effects we obtained reflect the processing and integration of human action sound features into distinctive representations, which may relate to a mechanism toward attaining a sense of meaningfulness to the listener. Compared to non-human sounds, the auditory information related to human actions contains many correlated features, some of which are highly salient to the infant (e.g., features related to animacy, meaningful action events, associated visual information, etc.) (Rakison, 2010; Rakison and Poulin-Dubois, 2001). Integrating all these features across multiple exemplars into a coherent representation, as well as updating previous representations with the information provided by new exemplars, could lead to the larger frontal PSW. The results concerning human non-speech vocalizations showed only partially this effect. Although they elicited a frontal PSW similar to the other human produced sounds, this was only marginally different from that recorded for non-human sounds. One could argue that the observed PSW differences between human produced and non-living sounds are related to the fact that the stimuli were not equally attended. In infants, auditory stimuli which are novel and which have potential biological significance (e.g., a deviant sound in a stream of otherwise repetitive sounds; mother vs. stranger voice) elicit attentional responses as indexed by a larger Pc signal at frontal and central locations (Kushnerenko et al., 2013; Purhonen et al., 2005). Although we had screened and chosen stimuli presumed to have similar probability to be encountered by infants in their surrounding environment, one could argue that some of the sound categories included in this study are more frequently encountered and/or more salient for the infant than others. These properties would differently elicit infants' attention. For example, the human produced sounds could be more salient to the infant due to their social nature and thus could be attended more. On the opposite, one could also argue that the human action sounds are less frequent in the infants' environment and thus elicit increased attention due to their novelty. However, our results do not currently support such interpretations. The Pc recorded at the frontal location suggests that the stimuli elicited similar levels of attention orientation, and thus the observed differences seem more likely to be related to their specific acoustic features. Nevertheless, future studies should specifically manipulate the degree of familiarity for these types of sounds, in order to more accurately establish whether they elicit increased allocation of attention due to their novelty or social saliency (de Haan and Nelson, 1997; Quinn et al., 2006; Kooijman et al., 2005).

The auditory stimuli used in this study were chosen for their ecological validity, thus preserving some of their natural psychophysical properties. By the age of 7-months, infants manifest sensitivity to many properties of sounds like intensity and frequency (e.g., Berg and Boswell, 1998, 1999), temporal relations (e.g., Baruch et al., 2004; Nazzi et al., 1998), and melody (e.g., Plantinga and Trainor, 2005; Trainor and Trehub, 1992). Importantly, the prenatal and postnatal development of the auditory system, as well as how infants attend to the auditory world around them, is strongly linked to these abilities (Burnham and Mattock, 2010). The present study newly revealed that category-specific conceptual processing observed in adults may already be present in incipient forms as early as 7-months after birth. The sound stimuli we used were carefully matched for acoustic features, therefore psychophysical properties cannot account for the obtained results. Nevertheless, it will be particularly interesting for future research to show which, if any, of these properties may be driving the responses to biological events, thereby revealing fundamental organizational principles during neurodevelopment.

In adults, an extended brain network has been functionally linked to the processing of auditory information about human actions, which partially overlaps with the one associated with processing related visual information (Engel et al., 2009; Lewis et al., 2011; Kable et al., 2005; Calvert et al., 2000). Auditory cues about action and motion are critical for our ability to detect others in our environment and to understand the meaning of their actions (Thomas and Shiffrar, 2010). Our study contributes to this line of investigation by providing the first evidence for a specific pattern of ERP responses to human action sounds in 7-month-old infants. These responses might reflect the activity of an extended brain network that specializes for processing information related to human action. Current results converge with a large body of literature showing significant developments in the ability to process action-related information after the age of 6-months (e.g., Daum et al., 2008; Woodward, 2013), as well as increasing cortical specialization for processing different types of social information around the same age (e.g., Blasi et al., 2011; de Haan et al., 2003; Pascalis et al., 2002). One potential direction for future research could be directed at understanding the ontogeny of this pattern of brain responses to human action sounds or, more generally, to sounds associated with 'living things'. It will be very interesting to find out whether the observed anterior temporal dissociation between human action sounds and the other categories of sounds is present at an earlier age and whether the parietal and frontal differentiation between 'living' versus

'non-living' sounds becomes more distinct in older infants. Of particular importance would also be to verify whether the pattern of ERP responses we report is restricted to human action sounds. A large body of literature shows that processing social stimuli like faces and voices goes through a process of perceptual narrowing (see Maurer and Werker, 2014 for a review). Thus it will be important to investigate whether processing action sounds is specific to human actions. All these potential lines of work will provide both confirmation of our current results and new knowledge. Moreover, future research should aim at understanding the role of human action sounds in the development of infants' abilities to understand others' actions, as well as its potential value as a predictive marker task for the early identification of atypical developmental outcomes.

Acknowledgements

References

Alaerts, K., Swinnen, S.P., Wenderoth, N., 2011. Action perception in individuals with congenital blindness or deafness: how does the loss of a sensory modality from birth affect perception-induced motor facilitation? J. Cogn. Neurosci. 23, 1080–1087.

Baruch, C., Panissal-Vieu, N., Drake, C., 2004. Preferred perceptual tempo for sound sequences: comparison of adults, children, and infants. Percept. Motor Skills 98, 325–339.

Belin, P., Fecteau, S., Bédard, C., 2004. Thinking the voice: neural correlates of voice perception. Trends Cogn. Sci. 8, 129–135.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 403, 309–312.

Benavides-Varela, S., Gómez, D.M., Mehler, J., 2011. Studying neonates' language and memory capacities with functional near-infrared spectroscopy. Front. Psychol. 2, 64, http://dx.doi.org/10.3389/fpsyg.2011.00064.

Berg, K.M., Boswell, A.E., 1998. Infants' detection of increments in low-and high-frequency noise. Percept. Psychophys. 60, 1044–1051.

Berg, K.M., Boswell, A.E., 1999. Effect of masker level on infants' detection of tones in noise. Percept. Psychophys. 61, 80–86.

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., et al., 2011. Early specialization for voice and emotion processing in the infant brain. Curr. Biol. 21, 1220–1224.

Brand, R.J., Tapscott, S., 2007. Acoustic packaging of action sequences by infants. Infancy 11, 321–332.

Burnham, D., Mattock, K., 2010. Auditory development. Wiley-Blackwell Handbook Infant Dev. 1, 81–119.

Calvert, G.A., Campbell, R., Brammer, M.J., 2000. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Curr. Biol. 10, 649–657.

Cheng, Y., Lee, S.Y., Chen, H.Y., Wang, P.Y., Decety, J., 2012. Voice and emotion processing in the human neonatal brain. J. Cogn. Neurosci. 24, 1411–1419.

De Haan, M. (Ed.), 2007. Infant EEG and Event-Related Potentials. Psychology Press.

Daum, M.M., Zmyj, N., Aschersleben, G., 2008. Early ontogeny of action perception and production. In: Morganti, F., Carassa, A., Riva, G. (Eds.), Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interactions. IOS Press, Amsterdam, The Netherlands, pp. 175–186.

Dehaene-Lambertz, G., Dehaene, S., Hertz-Pannier, L., 2002. Functional neuroimaging of speech perception in infants. Science 298, http://dx.doi.org/10.1126/science.1077066.

de Haan, M., Johnson, M.H., Halit, H., 2003. Development of face-sensitive event-related potentials during infancy: a review. Int. J. Psychophysiol. 51, 45–58.

de Haan, M., Nelson, C.A., 1997. Recognition of the mother's face by six-month-old infants: a neurobehavioral study. Child Dev. 68, 187–210.

Engel, L.R., Frum, C., Puce, A., Walker, N.A., Lewis, J.W., 2009. Different categories of living and non-living sound-sources activate distinct cortical networks. Neuroimage 47, 1778–1791.

Fecteau, S., Armony, J.L., Joanette, Y., Belin, P., 2004. Is voice processing species-specific in human auditory cortex? An fMRI study. Neuroimage 23, 840–848.

Friederici, A.D., Friedrich, M., Christophe, A., 2007. Brain responses in 4-month-old infants are already language specific. Curr. Biol. 17, 1208–1211.

Glushko, R.J., Maglio, P.P., Matlock, T., Barsalou, L.W., 2008. Categorization in the wild. Trends in Cognitive Sciences 12 (4), 129–135.

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Alsop, D., Detre, J., Gee, J., 2002. The neural basis for category-specific knowledge: an fMRI study. NeuroImage 15, 936–948.

Grossmann, T., Johnson, M.H., 2007. The development of the social brain in human infancy. Eur. J. Neurosci. 25, 909–919.

Grossmann, T., Striano, T., Friederici, A.D., 2005. Infants' electric brain responses to emotional prosody. NeuroReport 16, 1825–1828.

Grossmann, T., Oberecker, R., Koch, S.P., Friederici, A.D., 2010. The developmental origins of voice processing in the human brain. Neuron 65, 852–858.

Guillem, F., Bicu, M., Debruille, J.B., 2001. Dissociating memory processes involved in direct and indirect tests with ERPs to unfamiliar faces. Cogn. Brain Res. 11, 113–125.

Hillis, A.E., Caramazza, A., 1991. Category specific naming and comprehension impairment: A double dissociation. Brain 114, 2081–2094.

Hirai, M., Fukushima, H., Hiraki, K., 2003. An event-related potentials study of biological motion perception in humans. Neurosci. Lett. 344, 41–44.

Hirai, M., Hiraki, K., 2005. An event-related potentials study of biological motion perception in human infants. Cogn. Brain Res. 22, 301–304.

Kable, J.W., Kan, I., Wilson, A., Thompson-Schill, S., Chatterjee, A., 2005. Conceptual representations of action in lateral temporal cortex. J. Cogn. Neurosci. 17, 1855–1870.

Keysers, C., Kohler, E., Umiltà, M.A., Nanetti, L., Fogassi, L., Gallese, V., 2003. Audiovisual mirror neurons and action recognition. Exp. Brain Res. 153, 628–636.

Kooijman, V., Hagoort, P., Cutler, A., 2005. Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. Cogn. Brain Res. 24, 109–116.

Kuhl, P.K., 2010. Brain mechanisms in early language acquisition. Neuron 67, 713–727.

Kusak, G., Grune, K., Hagendorf, H., Metz, A.M., 2000. Updating of working memory in a running memory task: an event-related potential study. Int. J. Psychophysiol. 39, 51–65.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., Huotilainen, M., Näätänen, R., 2002. Maturation of the auditory event-related potentials during the 1st year of life. Neuroreport 13, 47–51.

Kushnerenko, E.V., Van den Bergh, B.R., Winkler, I., 2013. Separating acoustic deviance from novelty during the first year of life: a review of event-related potential evidence. Front. Psychol. 4.

Lewis, J.W., Brefczynski, J.A., Phinney, R.E., Janik, J.J., DeYoe, E.A., 2005. Distinct cortical pathways for processing tool versus animal sounds. J. Neurosci. 25, 5148–5158.

Lewis, J.W., Phinney, R.E., Brefczynski-Lewis, J.A., DeYoe, E.A., 2006. Lefties get it right when hearing tool sounds. J. Cogn. Neurosci. 18, 1314–1330.

Lewis, J.W., Talkington, W.J., Walker, N.A., Spirou, G.A., Jajosky, A., Frum, C., Brefczynski-Lewis, J.A., 2009. Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. J. Neurosci. 29, 2283–2296.

Lewis, J.W., Talkington, W.J., Puce, A., Engel, L.R., Frum, C., 2011. Cortical networks representing object categories and high-level attributes of familiar real-world action sounds. J. Cogn. Neurosci. 23, 2079–2101.

Lewis, J.W., Talkington, W.J., Tallaksen, K.C., Frum, C.A., 2012. Auditory object salience: human cortical processing of non-biological action sounds and their acoustic signal attributes. Front. Syst. Neurosci., http://dx.doi.org/10.3389/fnsys.2012.00027.

Lloyd-Fox, S., Blasi, A., Everdell, N., Elwell, C.E., Johnson, M.H., 2011. Selective cortical mapping of biological motion

processing in young infants. J. Cogn. Neurosci. 23, 2521–2532, http://dx.doi.org/10.1162/jocn.2010.21598.

Marshall, P.J., Shipley, T.F., 2009. Event-related potentials to point-light displays of human actions in 5-month-old infants. Dev. Neuropsychol. 34, 368–377.

Maurer, D., Werker, J.F., 2014. Perceptual narrowing during infancy: a comparison of language and 619 faces. Dev. Psychobiol. 56 (2), 154–178.

Moon, C., Fifer, W., 1986. New-born infants prefer the sound of mother's voice as experienced in the womb. Paper presented at the International Society for Developmental Psychobiology Annual Meeting, Annapolis, Maryland, November.

Moore, C.J., Price, C.J., 1999. A functional neuroimaging study of the variables that generate category-specific object processing differences. Brain 122 (5), 943–962.

Morrongiello, B.A., Clifton, R.K., 1984. Effects of sound frequency on behavioral and cardiac orienting in newborn and five-month-old infants. J. Exp. Child Psychol. 38, 429–446.

Nazzi, T., Bertoncini, J., Mehler, J., 1998. Language discrimination by newborns: toward an understanding of the role of rhythm. J. Exp. Psychol.: Human Percept. Performance 24, 756.

Parise, E., Friederici, A.D., Striano, T., 2010. Did you call me? 5-Month-old infants own name guides their attention. PloS One 5, e14208.

Pascalis, O., de Haan, M., Nelson, C.A., 2002. Is face processing species-specific during the first year of life? Science 296, 1321–1323.

Pena, M., Maki, A., Kovacic, D., Dehaene-Lambertz, G., Koizumi, H., Bouquet, F., Mehler, J., 2003. Sounds and silence: an optical topography study of language recognition at birth. Proc. Natl. Acad. Sci. USA 100, 11702–11705.

Peña, M., Pittaluga, E., Mehler, J., 2010. Language acquisition in premature and full-term infants. Proc. Natl. Acad. Sci. 107 (8), 3823–3828.

Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., Di Russo, F., 2005. Separate neural systems for processing action-or non-action-related sounds. Neuroimage 24, 852–861.

Plantinga, J., Trainor, L.J., 2005. Memory for melody: infants use a relative pitch code. Cognition 98, 1–11.

Purhonen, M., Kilpeläinen-Lees, R., Valkonen-Korhonen, M., Karhu, J., Lehtonen, J., 2005. Four-month-old infants process own mother's voice faster than unfamiliar voices—electrical signs of sensitization in infant brain. Cogn. Brain Res. 24, 627–633.

Rakison, D.H., 2010. Perceptual Categorization and Concepts. Wiley-Blackwell Handbook of Infant Development, vol. 1, second ed., pp. 243–270.

Rakison, D.H., Poulin-Dubois, D., 2001. Developmental origin of the animate–inanimate distinction. Psychol. Bull. 127, 209–228.

Reddy, R.K., Ramachandra, V., Kumar, N., Singh, N.C., 2009. Categorization of environmental sounds. Biol. Cybernet. 100, 299–306.

Quinn, P.C., Westerlund, A., Nelson, C.A., 2006. Neural markers of categorization in 6-month-old infants. Psychol. Sci. 17, 59–66.

Sievers, B., Polansky, L., Casey, M., Wheatley, T., 2013. Music and movement share a dynamic structure that supports universal expressions of emotion. Proc. Natl. Acad. Sci. 110, 70–75.

Silveri, M.C., Gainotti, G., Perani, D., Cappelletti, J.Y., Carbone, G., Fazio, F., 1997. Naming deficit for non-living items: neuropsychological and PET study. Neuropsychologia 35, 359–367.

Tchernichovski, O., Mitra, P.P., Lints, T., Nottebohm, F., 2001. Dynamics of the vocal imitation process: how a zebra finch learns its song. Science 291, 2564–2569.

Thomas, J.P., Shiffrar, M., 2010. I can see you better if I can hear you coming: action-consistent sounds facilitate the visual detection of human gait. J. Vision 10, 14.

Trainor, L.J., Trehub, S.E., 1992. A comparison of infants' and adults' sensitivity to Western musical structure. J. Exp. Psychol.: Human Percept. Perform. 18, 394–402.

Trainor, L.J., 2010. Using electroencephalography (EEG) to measure maturation of auditory cortex in infants: Processing pitch duration and sound location. In: Tremblay, R.E., Barr, R.G., Peters, R.deV., Boivin, M. (Eds.), Encyclopedia on Early Childhood Development. Centre of Excellence for Early Childhood Development, Montreal, Quebec, pp. 1–5.

Van der Zwan, R., MacHatch, C., Kozlowski, D., Troje, N.F., Blanke, O., Brooks, A., 2009. Gender bending: auditory cues affect visual judgements of gender in biological motion displays. Exp. Brain Res. 198, 373–382.

Warrington, E.K., Shallice, T., 1984. Category specific semantic impairments. Brain 107, 829–854.

Wilding, E.L., Doyle, M.C., Rugg, M.D., 1995. Recognition memory with and without retrieval of context: an event-related potential study. Neuropsychologia 33, 743–767.

Winkler, I., Kushnerenko, E., Horváth, J., Ceponiene, R., Fellman, V., Huotilainen, M., Sussman, E., 2003. Newborn infants can organize the auditory world. Proc. Natl. Acad. Sci. USA 100, 11812–11815.

Woodward, A.L., 2013. Infant foundations of intentional understanding. In: Banaji, M.R., Gelman, S.A. (Eds.), Navigating the Social World: A Developmental Perspective. Oxford University Press, Oxford, pp. 75–80.