Gaze direction and maternal pitch in surprise-eliciting situations

Nadja Reissland*, John Shepherd

Department of Psychology, William Guild Building, University of Aberdeen, Aberdeen AB24 2UB, Scotland, UK

Received 30 July 2001; received in revised form 18 December 2001; accepted 30 December 2001

Abstract

The purpose of this study was to test the hypothesis that not only do babies use emotional signals from adults in order to relate emotions to specific situations (e.g., Campos & Stenberg, 1981) but also that mothers seek out emotional information from their infants (Emde, 1992). Three groups of mothers and their infants, 3, 5 and 9 months old were video- and audio-taped, while playing in their homes with a soft toy and a remote-control Jack-in-the-box. During surprise-eliciting play with the Jack-in-the-box, maternal and infant gaze direction and their emotional expressions of surprise, pleasure, fear and neutral expressions were coded in three regions of the face. In addition, the mean fundamental frequency of maternal surprise-vocalisations was analysed. Maternal exclamations of surprise were compared with similar utterances of these mothers while playing with a soft toy as a baseline. During the surprise event, maternal and infant gaze directions as well as infant age were analysed in relation to maternal pitch. Results are discussed in terms of maternal use of the pitch of her voice to mark surprising situations, depending on the gaze-direction of the infant. © 2002 Published by Elsevier Science Inc.

Keywords: Gaze direction; Pitch of surprise; Emotional-development

1. Introduction

Research on referencing observes how babies use emotional signals from adults in order to relate emotions to specific situations (e.g., Campos & Stenberg, 1981; Feinman & Lewis, 1983; Camras & Sachs, 1991; Gunnar & Stone, 1984; Klinnert, Emde, Butterfield, & Campos, 1986; Mumme, Fernald, & Herrera, 1996). According to Feinman (1991), adults can guide
infant learning by interpreting for the infant the meaning of a situation, by for example smiling and touching a toy which the infant seems puzzled over. Infants, in such puzzling situations will look at the mother for guidance on how to behave (Campos & Stenberg, 1981).

These studies on social referencing have typically focussed on the infant, rather than the adult, receiving and interpreting information about the stimulus. However, as Emde (1992, p. 82) pointed out, “Mothers seek out” emotional information from their infants in order to guide caregiver action”. Hence, in the construction of social reality both mother and infant reference from each other. The focus of the present research was the mother.

Infants can express emotions, for example pleasure, from birth (e.g., Wolff, 1987; Sroufe, 1996). Some people would extent this to include most emotions (Izard, 1971). In the case of surprise, however, it has been argued that facial expressions of the emotion emerge only after 8 months, when the infant has developed the concept of object permanence and can distinguish between expected and unexpected outcomes (Charlsworth, 1969). While emotional facial expressions may be innate, the conditions under which to show appropriate expressions of emotion are learned in the process of development (Camras, 1994; Charlsworth, 1969; Fridlund, 1994; Sroufe, 1996).

Emotional reactions in social referencing situations have been usually related to facial expressions (e.g., Campos & Stenberg, 1981; Feinman & Lewis, 1983; Camras & Sachs, 1991; Gunnar & Stone, 1984; Klinnert, Campos, Sorce, Emde, & Svejda, 1983; Klinnert et al., 1986). However, children can also be guided by paralinguistic aspects of emotional expressions, such as pitch of voice (e.g., Fernald, 1993; Mumme et al., 1996; Walker-Andrews, 1997; Walker-Andrews & Lennon, 1991). Depending on the situation, mothers use different paralinguistic cues (Fernald, 1989; Reissland & Snow, 1996; Reissland, 2000). Mothers differentiate between soothing or arousing contexts by using different pitch (e.g., Papousek, 1992; Reissland, 2000) and they differentiate in their pitch between play and non-play (Reissland & Snow, 1996), as well as between praising and scolding utterances (Fernald, 1992).

Preverbal infants respond to vocal expressions of emotion as well as facial expressions (Mumme et al., 1996; Wolff, 1987). For example, infants smile consistently to infant-directed speech when presented alone, but not when a face is presented by itself (Wolff, 1987). Walker (1982) found that when 7-months old infants saw side by side dynamic facial displays of happiness and anger and heard a vocal expression corresponding to either happiness or anger, that these infants preferred to look at a face displaying the emotion corresponding to the vocal expressions. Soken and Pick (1999) extended this study to two positive and two negative emotions. They found that 7 months olds distinguished happy, interested, angry and sad expressions, looking longer when they heard concordant vocal and facial displays than at the dissonant displays. Hence, infants are sensitive not only to the emotion shown on the face but also to the specific vocal information of the emotional expression. With regard to social referencing, Mumme et al. (1996) suggest that even at 12 months, infants rely on vocal expressions of fear when referencing their mothers’ emotional expressions. Therefore, one important function of intonation in speech to preverbal infants is the communication of affect (Fernald, 1989; Papousek & Papousek, 1981; Stern, 1985).

Various emotions can be differentiated on the basis of vocal parameters. Scherer (1982a, 1982b) reported that exclamations of surprise are distinguished from vocalisations of other emotions such as anger, fear, shame or pleasure by auditory parameters. A high pitch level
(measured in Hz) is strongly associated with surprise (Scherer, 1982a). Although pitch is important in emotional expressions, it is not known whether mothers adjust the pitch of their vocal expressions of surprise depending on their child’s gaze direction, or whether maternal pitch height is independent of the child’s gaze direction and facial expression.

Following Charlsworth (1969) and Camras (1994), it is expected that younger infants in the study will show less surprise in the context of the surprise-eliciting situation in comparison with the oldest group. In the surprise-eliciting situation mothers will monitor their infants emotional expression and will react by varying their vocal pitch depending on whether the infant is looking at her facial expressions or the Jack-in-the-box.

2. Method

2.1. Participants

Three groups of Scottish mothers and their infants (22 infants, ranging from 2 to 4 months, mean age = 3.5 months; 22 infants, ranging from 5 to 7 months, mean age = 5.7 months; and 22 infants, ranging from 8 to 10 months, mean age = 9.4 months) were video- and audio-taped in their homes. Their mothers were monolingual English speakers and were primary caregivers. Mothers had a mean number of 14 years in education, ranging from 10 to 19 years.

2.2. Procedure

Infants and their mothers were observed while playing with a remote controlled Jack-in-the-box placed at an angle of 45° to the midline of mother and child, facing each other and positioned at or below the child’s line of gaze. The Jack-in-the-box is commercially not available and hence mothers and infants were exposed to a novel stimulus. Mother and infant also played with a soft toy during the session. Two video cameras were used, one focusing on the mother and the other focusing on the child. Infants sat in a baby-chair or propped up with a pillow. Mothers, facing their infants, were instructed to open the Jack-in-the-box by pushing a button, which activated a mechanism to open the lid of the Jack-in-the-box. Mothers were given no direction as to their own expressions of emotions.

2.3. Acoustic measurements

For each mother the first vocal exclamation of surprise (e.g., “oh”) immediately following the opening of the lid of the Jack-in-the-box and which did not contain background noise, was recorded. The first spontaneous exclamation of surprise was selected because none of the mothers had previously seen the particular Jack-in-the-box used in this study, which is commercially not available, and the first exclamation was deemed to be the most “natural” expression of surprise by the mother. Sometimes, the first surprise exclamation could not be used because of extraneous noise (e.g., moving of the tripod) and then the next exclamation was used. For the baseline-recording the first maternal utterance of “oh” during play with the
soft toy, which was considered a pleasurable situation, was recorded. Only vocalisations free from background noise were analysed.

All vocalisations selected for analysis were digitised using Signalise (Keller, 1994). This package is designed to analyse the acoustic aspects of speech. The pitch extraction algorithm (Fast Fourier transform) of this programme reports fundamental frequency values \( (F_0) \) for each 5 ms of voiced speech. For each utterance, the following measures were obtained: minimum, maximum, and mean \( F_0 \) values measured in Hz. The amplitude was measured by a pitch-synchronous energy calculation. Minimum, maximum and mean amplitude, measured in dB was obtained for each utterance. The length of each utterance was measured in seconds.

2.4. Maternal and Infant facial expression

Facial expressions of pleasure, fear, neutral expression or surprise were coded from three regions of the face, adapted from a coding scheme described by Hiatt, Campos, and Emde (1979) and Stenberg, Campos, and Emde (1983) and corresponding to AC (appearance change) codes developed by Izard (Max coding system, 1979, 1995). Surprise is recorded in the eyebrow/forehead region, if brows are raised in the normal contour but they are curved and high and the skin below the brow is stretched and/or horizontal wrinkles are visible across the forehead. This corresponds to AC 20 in the Max coding system. Surprise is shown in the eyes/lid region, if the eyelids are open, the upper lid is raised and the lower lid is drawn down, corresponding to AC 30 in the Max coding system. If the jaw drops open so the lips and teeth are parted and but there is no tension or stretching of the mouth surprise is recorded in the nose/mouth region, corresponding to AC 50 in the Max coding system. Fear is coded in the eyebrow/forehead region, if the brows are lowered and drawn together and/or the wrinkles of the forehead are across the centre of the forehead and not the entire forehead or there is thickening in the mid-region of the forehead. This corresponds to AC 22 in the Max coding system. Fear of the nose/mouth region is coded if the mouth is open and the lips are tensed and drawn back or stretched and drawn back, corresponding to AC 53 in the Max coding system. Pleasure is coded for the eyes/lid region if the upper eyelid corners are raised, and crows feet wrinkles go outward from the corners of the eyes. This corresponds to AC 33, which is described as a squint. Pleasure is credited for the nose/mouth region if the corners of the lips are drawn back and up and teeth may or may not be visible. This corresponds to AC 52 in the MAX coding system. Neutral facial expressions were coded for all three regions of the face if there was no visible movement in these regions of the face. Mother and infant were credited with showing surprise, pleasure, fear, or neutral expressions in no region (=0) or 1–3 regions of the face.

2.5. Reliability

Reliability of observational data was obtained on the basis of all codes used from the time the lid of the Jack-in-the-box opened to the time it was closed and calculated with Cohen’s Kappa for 15% of the infants and 15% of the mothers. For the infants, Cohen’s Kappa ranged from .719 to .918. For the mothers, Cohen’s Kappa ranged from .736 to .791.
3. Results

3.1. Vocal aspects of surprise exclamations in relation to pleasure: baseline measurement

A paired $t$-test of the mean fundamental frequency of maternal surprise (mean $F_0 = 377.32$ Hz) and pleasure exclamations (mean $F_0 = 290.51$ Hz) was significant ($t = 10.2$, $df = 65$, $p < .001$, two-tailed). Hence, mothers used a higher mean fundamental frequency when exclaiming in surprise in comparison with similar exclamations of pleasure (see Table 1).

3.2. Expressions of surprise, pleasure, fear and neutral facial expressions by infants in relation to the Jack-in-the-box

A MANOVA on the expression of each emotion by the infants (surprise while lid of Jack-in-the-box opened until mother exclaimed in surprise, surprise while mother exclaimed in surprise until lid was closed, pleasure, fear and neutral facial expressions during the whole episode) was carried out with age group as the independent variable. Expressions were coded across three regions of the face for the infants. The codes ranged from 0 to 3 regions of the face showing the four emotions.

Analysing the number of regions (0–3) of the face showing the four emotions, there was no significant age difference for the expression of fear ($F(2, 63) = .069, p > .933$), pleasure ($F(2, 63) = .980, p > .381$) or neutral expressions ($F(2, 63) = .007, p > .926$) among the infants.

There was a significant effect for infants showing surprise facial expression in 0–3 regions of the face during the popping up of the Jack-in-the-box ($F(2, 63) = 5.56, p < .006$). A pairwise comparison with Bonferroni adjustment for multiple comparisons, established that 3.5 months old infants showed significantly fewer elements of surprise (mean .36 regions of the face) in comparison with 9.4 months old infants (mean 1.50 regions of the face). There was a significant effect for infants showing surprise facial expression in 0–3 regions of the face during maternal exclamations of surprise ($F(2, 63) = 26.01, p < .001$). A pairwise comparison with Bonferroni adjustment for multiple comparisons, established that 3.5 months old infants showed significantly fewer elements of surprise (mean .59 regions of the face) in comparison with 5.7 months old infants (mean 1.50 regions of the face) and 9.4 months old infants (mean 2.4 regions of the face, see Table 1).

3.3. Maternal facial expressions of surprise, pleasure, fear and neutral in relation to the operation of the Jack-in-the-box

A MANOVA on the expression of each emotion by the mothers (surprise while lid of Jack-in-the-box opened until mother exclaimed in surprise, surprise while mother exclaimed in surprise until lid was closed, pleasure, fear and neutral facial expressions) was carried out with age group of infant as the independent variable. Expressions were coded across three regions of the face for the mothers. The codes ranged from 0 to 3 regions of the face showing the four emotions.

There was no significant effect for mothers showing fear ($F(2, 63) = .936, p > .398$), pleasure ($F(2, 63) = .324, p > .725$) and neutral facial expressions ($F(2, 63) = .870, p > .424$),
Table 1

Pitch of maternal vocal expressions of pleasure and surprise and mean number of regions of mothers’ and children’s faces showing happiness, surprise, fear or neutral facial expressions

<table>
<thead>
<tr>
<th>Age groups of children (months)</th>
<th>Mean Hz of mother’s voice</th>
<th>Mean no. of regions of mothers (m) and children (c) showing emotional expression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Happy</td>
<td>Surprise (m to noise of Jack&lt;sup&gt;1&lt;/sup&gt;, while vocalising surprise&lt;sup&gt;2&lt;/sup&gt;)</td>
</tr>
<tr>
<td>2–4</td>
<td>293.99 (52.23)</td>
<td>409.58 (78.71)</td>
</tr>
<tr>
<td>5–7</td>
<td>294.01 (34.38)</td>
<td>368.88 (73.17)</td>
</tr>
<tr>
<td>8–10</td>
<td>283.52 (29.05)</td>
<td>353.49 (51.59)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Standard deviation are in parentheses.
from the time the lid of the Jack-in-the-box opened until the lid of the Jack-in-the-box was closed in relation to the three age groups of infants. There was no significant effect for mothers showing surprise while the lid of the Jack-in-the-box opened till the mother started to exclaim in surprise, \( F(2, 63) = 1.791, p < .175 \).

There was, however, a significant effect of maternal surprise expressions while she exclaimed in surprise until the lid was closed \( F(2, 63) = 4.92, p < .01 \). A pairwise comparison with Bonferonni adjustment for multiple comparisons, established that mothers of 5.7 months old infants showed significantly fewer elements of surprise (mean 1.91 regions of the face) in comparison with 9.4 months old infants (mean 2.73 regions of the face, see Table 1).

### 3.4. Direction of gaze by mother and infant

In an attempt to identify the source of referencing during vocal and facial expressions of surprise, infant and maternal gaze directions were identified as the Jack-in-the-box opened. The Jack-in-the-box was placed at an angle of 45° to the midline of mother and infant, so that one could clearly discriminate between the infant’s gaze direction, at mother or Jack-in-the-box. A sign test \( z = 4.63, p < .001 \) established that significantly more mothers gazed at their infants’ face when the box opened (82%) than infants gazed at their mothers’ face (36%). There were age differences among infants, with more infants in the youngest group gazing at their mother (64%) in comparison with the infants of the middle and oldest group (23%) \( \chi^2 = 4.45, df = 1, p < .03 \). Hence, the youngest group of infants looked at the mother whereas the older infants looked at the Jack-in-the-box.

In order to identify whether looking at the mother had an effect on children’s facial expressions of emotions, children were tested in respect to elements of surprise shown while looking at the mother exclaiming in surprise. An ANOVA for the whole sample with gazing (at mother or not gazing at mother) as independent variable and with elements of surprise (0–3) shown as dependent variable, was carried out. There was no effect of gazing at mother \( F(2, 63) = .279, p = .59 \). Hence, gazing at mother while she expressed surprise vocally did not affect the number of elements of surprise shown in the babies face, despite the fact that mothers did show exaggerated surprise expressions when exclaiming with surprise (see Table 1).

### 3.5. The pitch of mother’s voice and gaze direction of the infant

In order to identify whether maternal pitch height differences were due to infants’ gaze direction or infants’ age a two-way ANOVA on the pitch of maternal surprise exclamations was carried out with age group and gaze direction of child as the independent variables. There was a significant main effect for gaze direction of child \( F(1, 60) = 4.831, p < .032 \) with the pitch being higher for children looking at the mother than for children not looking at their mother, but neither the age of child main effect \( F(2, 60) = 2.237, ns \) nor the age group by gaze direction interaction was significant \( F(2, 60) = .956, ns \). This indicates that the differences in maternal pitch were due to the infant’s gaze direction rather than the infant’s age.

Conceivably, mothers pitch might govern infant gaze direction. In order to test this possibility only those infants who looked at the mother while she exclaimed in surprise were tested.
If infants were drawn to look at the mother because of the height of her surprise exclamation then the pitch of maternal exclamations should be higher for those infants who looked at her while or just after she had exclaimed in surprise, compared to infants who had started to look at her before she exclaimed in surprise. In order to test this hypothesis, infants looking at their mother before she had exclaimed in surprise were compared with infants looking at her after she had exclaimed in surprise. A t-test comparing the pitch of maternal surprise exclamation of mothers whose children looked at the mother before she exclaimed in surprise (mean $F_0 = 424.23$ Hz) with the pitch of maternal surprise exclamation of mothers whose children looked at the mother after she exclaimed in surprise (mean $F_0 = 391.16$ Hz, $t = .808$, df = 31, $p = .425$, two-tailed, ns) was non significant. Hence, the pitch of maternal surprise exclamations did not attract the child’s gaze.

4. Discussion

Social referencing (Campos & Stenberg, 1981) is a communicative process in which both partners react to each other (Emde, 1992). In the present study, the focus was on the mother. The pitch of the mothers’ voice was related to the gaze direction of the child. Mothers whose child looked at her while the Jack-in-the-box opened exclaimed with a higher pitch (mean: 410.82 Hz) in comparison with mothers whose child looked at the Jack-in-the-box (mean pitch: 358.17 Hz). Furthermore, more of the younger infants (64%) than the older infants (23%) looked at their mothers. The youngest infants showed significantly fewer elements of surprise than the oldest group of infants. These results suggest that mothers do not simply communicate affect depending on the situation, rather they adjust their emotional communication, specifically the pitch of their voice, depending on whether the infant looks at her or at the stimulus.

Campos, Mumme, Kermoian, and Campos (1994) suggested that in some contexts, emotional expressions are used as “social signals” in order to influence the behaviour of others. Social signals, specifically maternal expressions of emotion, can influence infants in their reactions to environmental stimuli. For example, Klinnert et al. (1983) found that depending on their mother’s facial expressions, infants would evaluate a toy as pleasant or unpleasant or the visual cliff as dangerous or safe to cross. In contrast, in the present study, where the focus was maternal behaviour, virtually all mothers looked at their infants when exclaiming in surprise, but only a minority of infants looked at their mothers. However, even for those children who gazed at their mother’s face, looking at the mother’s facial expression of surprise did not result in the child showing surprise in more regions of the face. Therefore, it seems that mothers exaggerate or attenuate their vocal expression of surprise while monitoring their infants’ direction of gaze rather than infant facial expression.

Although, according to Izard (1991), emotional expressions are innate and the facial expression and feeling state of the emotion correspond at least during the first year of life, Camras (1994) maintains that there is no complete concordance between surprise and situations eliciting surprise. The question whether facial expression and feeling state of emotion correspond cannot be addressed in this study. Nevertheless, according to the results of this present study, it appears that whether or not the child does feel surprised when it shows components of surprise is irrelevant for mothers. Mothers seem not to react to the infant facial expression of surprise
but whether or not the infant looks at her and hence are attentive to her. If children learn to express emotions in a culturally accepted way, mothers might assume that their children feel what they themselves feel in surprise eliciting situations, namely surprised. Sroufe (1996) argued that emotions are developmental constructs, in which meaning of the context plays an important role in experiencing an emotion. According to Sroufe (1996) if there is a lack of “subject–object” relationship then the emotional responses are reflexive. Hence, even though a new-born infant might show the facial configurations of ‘fear’ or ‘surprise’ this does not mean that the infant feels ‘fear’ or ‘surprise’. The results of the present study would suggest that part of the process to establish the meaning of the situation is the mother’s vocal response to that situation in which she monitors the child’s gaze direction. This research should be extended to other populations. For example for a blind mother the trigger to change her voice could be the touch rather than gaze direction of her baby. This would indicate that mothers react sensitively to what their infants are interested in and they make use of this interest in order to teach the child about emotional reactions.

Acknowledgments

We thank two anonymous reviewers for their constructive feedback, Mark Marshark, for his comments on a previous draft of the paper, Lily Cowie and Eisquel Herrera for collecting and coding part of the data, Ioanna Lipourli, Emma Mitchell, and Sarah Logan for coding part of the data as well as mothers and infants who took part in this study. The study was supported by a grant from the Leverhulme foundation ref. F/152/X to the first author.

References


