Talking with Feeling: Integrating Affective and Linguistic Expression in Early Language Development

Lois Bloom
Teachers College, Columbia University
Richard Beckwith
Princeton University


ABSTRACT

The purpose of the longitudinal study reported here was to determine the developmental relation between the two systems of expression available to the young child in the period of early language learning: affect and speech. Two achievements in language were identified for a group of 12 infants: First Words, at the beginning of the single-word period (mean age about 13 months), and a Vocabulary Spurt toward the end of the period (mean age about 19 months). Affect expression was coded continuously in the stream of the infants' activity as they and their mothers played with groups of toys and ate a snack. The occurrence of words was examined in relation to the expression of affect and the results of this study concern developments in the integration of these two forms of expression. The children's words occurred closely in time with their expression of emotionally toned affect. We concluded, then, that they were learning words to express what their feelings were about even though none of the actual words they said were emotion terms. However, the cognitive requirements for emotional expression and expression through speech resulted in several constraints on their integration. Words were said with neutral affect expression most often. Emotionally toned affect expression was least likely to occur before words, with a peak in emotional expression in the moments immediately after words. The peak in emotional expression with words was significantly greater, and the pre-word dip in emotional expression was significantly less, at the Vocabulary Spurt than at First Words. Thus, the two systems of expression converged in the period of single-word development as the children came to be able to say words with emotional affect. However, their emotional expression with words had positive rather than negative valence, with low rather than heightened intensity, and were among the children's most frequent and earliest learned words. These results are discussed in terms of the acquisition of language for expression and the different cognitive requirements for expression through affect and speech.

The first forms of affect expression are already available to infants in the beginning months of life and are probably biologically determined (e.g. Emde, Gaensbauer, & Harmon, 1976; Izard, 1977). By the time language begins in the second year, the development of the system of emotional expression is well underway. The purpose

---

1 The research reported here was supported by research grants from The National Science Foundation and The Spencer Foundation to Lois Bloom, and the James S. McDonnell Foundation to Princeton University. We thank Joanne Bitetti Capatides, Mariette Newton-Danilo, Erin Tinker, Anne Spangler, and Roxanne DeMatteo for their help with the analyses of data; Jeremie Hafitz, Matthew Rispoli, Joy Vaughn, Erin Tinker, Jennifer Saldano, and Scott Scheer for transcribing the children's speech; and Virginia Brennan, Suzanne Gottlieb, Mariette Newton-Danilo, Lisa Spiegel, and Pia Wikstrom for coding affect expression. We thank Robert Beckwith and Karen Clark for helping to develop the lag sequential analyses. We have benefited from discussions of the results of this study with Catherine Best, Joseph Jaffe, Margaret Lahey, Karin Lifter, Dennis Molfese, Nancy Stein, Don Tucker, and Gerald Turkewitz, in addition to the members of our research group.
of the study reported here was to determine how developments in emotional expression and language “fit” together in the course of early language development.

Developments in emotional expression consist of the infant gaining control over the mechanisms of expression, and modulating the forms of expression to conform to social and cultural expectations and practices (Izard & Malatesta, 1987). Language may be the most important context in which this learning occurs. First, language allows individuals to articulate something of the causes, the circumstances, and the objects of their emotional experience. The study reported here concerns the period of early word learning in the second year. This is the period when infants are learning words that can express the objects of their emotions or what their feelings are about, while they continue to express how and what they feel through facial and vocal affect displays (Bloom, Beckwith, Capatides, & Hafitz, 1988). Second, the linguistic resources that societies have evolved include words, structures, and discourse practices for expressing feelings and for modulating the facial and vocal expression of those feelings (Ochs & Schieffelin, 1989). However, the children in the present study had not yet begun to learn the words that name emotions or other forms of emotion language (addressed in other studies with somewhat older children, e.g. Bretherton, & Beeghly, 1982; Ridgeway, Waters, & Kuczaj, 1985).

In reviewing studies of vocalization in infants’ prespeech emotional expressions, Malatesta (1981, p. 15) suggested that “the tendency to intonate speech affectively may have an early and innate origin”. The nonspeech vocalizations of 8-month-old infants occur at times of obvious shifts in attention and changes in facial expression, particularly smiles (Hilke, 1988). We know that the words and sentences of school-aged children (e.g. Camras, 1985) and adults (e.g. Davitz 1964, 1969; Scherer, 1986) are often spoken with emotional expression. However, no studies, as far as we know, have examined relations between speech and affect expression in the period of language development in the second year, after the period of infant prespeech vocalization and before the acquisition of mature language.

COGNITION, LANGUAGE, AND EMOTION

Both speech and affect express something of the underlying contents of desires and beliefs, plans and goals. Expression, whether speech or affect, requires the allocation of cognitive resources for representing such contents of mind. However, the requirements for speech and affect expression also differ. Using language requires comparing the contents of attention with prior experience recalled from memory, recalling the linguistic forms to express these contents, and encoding (Bloom & Beckwith, 1988). The expression of emotion entails an evaluation (e.g. Frijda, 1986; Scherer, 1984; Wozniak, 1986) which includes, for example, considering the consequences of an aspect of the situation for some goal or plan (Oatley & Johnson-Laird, 1987; Stein & Levine, 1987), along with a subjective feeling and a display (e.g. Izard, 1977; Lewis & Michalson, 1985). For words to express something of what an emotional expression is about, cognitive resources need to be allocated for both speech and affect expression. Thus, the requirements for the two sorts of expression could compete with one another when the young child is expressing emotion and wants to say a word at the same time, and vice versa.

The competition for cognitive resources to meet the requirements for speech and affect expression may be most apparent in the early period of language learning, before the processes of lexical recall and speech become automatic. In previous research, we observed that 1-year-old infants who spent relatively more time in neutral affect expression achieved certain milestones in language learning earlier than infants who expressed emotionally toned affect more frequently (Bloom & Capatides, 1987a; Bloom, Beckwith, & Capatides, 1988). We have interpreted this result to mean that neutral affect expression allows the deployment of cognitive resources for early language learning. Learning to say words and expressing emotion may compete with one another for the young child’s attention and limited cognitive resources.

Such competition for cognitive resources may well have a neuropsychological basis related to what Kinsbourne has referred to as a “lateral gradient of attention.” When cognitive processing requires skills that are specialised for the different cerebral hemispheres, the result is a competition for “relative degree of activation” of the two hemispheres (Kinsbourne, 1975, p. 81). Emotional expression and language involve different regions of the brain. While the emotions involve both subcortical as well as cortical regions in the brain, emotional expression and monitoring emotional arousal have predominantly right hemisphere specialisation (e.g. see Best & Queen, 1989;
Evidence from infants suggests that cerebral asymmetries such as these are present in early infancy before the emergence of language, and may be predetermined rather than developmental (e.g. Best, 1988; Molfese, Freeman, & Palermo, 1975; Witelson, 1987). Integration of the processing activities in the two hemispheres and allocation of attentional capacities are mediated to a large degree by the corpus callosum (e.g. Levy, 1985). However, the corpus callosum is particularly slow to mature (Yakovlev & Lecours, 1967), and this fact has been cited to explain the relative lack of early hemispheric integration in other developmental domains, for example, graphomotor skill (Kirk, 1985) and face recognition (Levine, 1985).

In sum, different regions of the brain are specialized for language (the left hemisphere) and emotional expression (predominantly the right hemisphere). This separation of the neurological bases for the two kinds of expression, in different regions of the brain, could interfere with simultaneously saying a word and expressing emotion about the same content because of competition for activation between the two hemispheres. This constraint would limit the distribution of cognitive resources between the mental activities required for emotional expression and speech. In early language learning, in particular, expression through speech could consume the young child’s attention and interfere with the simultaneous expression of affect, and expression of emotion could preempt the attention required for speech. Although we expect children to learn language for expressing what their feelings are about, the two sorts of expression may be independent when language begins. However, with the automaticity of speech that comes as language develops, we should see their integration.

Three hypotheses were offered in the present study to test the assumption that a limitation in resources constrains saying words and expressing emotionally toned affect at the same time. The distribution of cognitive resources should be reflected in the temporal relation between saying a word and expressing emotion, if the word preempts the cognitive activity required for expressing emotion, and vice versa. The result would be non-overlap between the two forms of expression. The first hypothesis, the Integration hypothesis, was that words and affect expression would be independent initially but would begin to overlap as words increase in number and frequency, reflecting increased automaticity of use.

Another set of hypotheses concerned the factors that might contribute to the integration of words and affect in the single-word period in the second year. One such factor could be emotional valence and the difference between positive and negative affect expression. More cognitive work occurs with expression of the negative emotions than with positive emotion, because the positive and negative emotions differ with respect to the direction and the outcome of an evaluation of the situation relative to an individual’s goals and behavior (Oatley, 1988; Stein & Levine, 1989). With the positive emotions, a goal has been achieved and no new cognitive activity is required (Rothbart, 1973). In the case of the negative emotions, however, the result of evaluation includes cognitive activity for the construction of a new plan to remove an obstacle to the goal or the creation of a new goal (Stein & Jewett, 1987). Positive emotion is associated with the success of ongoing behavior whereas negative emotion is associated with the interruption of behavior (Kinsbourne, 1988). We might expect, then, that integrating affect expression and speech in the single word period would be more likely with positive than with negative affect expression.

Some evidence in the neuropsychological literature suggesting differences in regions of brain activity for the positive and negative emotions would support this expectation. While the neuropsychological evidence strongly supports a right-hemisphere specialization for emotional expression in general (e.g. Borod & Koff, 1989; Tucker, 1986), studies comparing brain-injured and normal populations indicate that certain aspects of the expression of positive emotions may involve left-hemisphere activity as well (e.g. Kinsbourne & Bemporad, 1984; Sackheim & Gur, 1978; Schwartz, Ahern, & Brown, 1979). EEG studies with infants have also suggested a valence asymmetry in specialization, with more left-hemisphere activity associated with positive emotional expressions and more right-hemisphere activity with negative expressions (e.g. Davidson & Fox, 1982). The issue of differential
hemisphere activation with positive and negative emotional expression is presently a controversial issue. However, if aspects of positive affect expression are attributable to left hemisphere activity, we might expect that speech would be more readily integrated with positive emotional expression than with negative emotional expression. Accordingly, the second hypothesis we tested was the Facilitating Affect hypothesis: Given the lesser cognitive load and the possibility of less interhemisphere competition with positive expression, the integration of speech and affect expression was expected to occur more often with positive than with negative expression.

Finally, saying words and simultaneously expressing emotion should not be independent of the words themselves. The words that children know best would presumably include their earlier learned and most frequent words. These should have greater automaticity and be more likely to occur together with emotional expression than words that the child is just learning. The third hypothesis we tested was the Facilitating Words hypothesis: Integration of speech and emotional expression was expected to occur with the children’s earlier learned and most frequent words since these are words they presumably know best.

In sum, the hypotheses tested in this study were that the integration of words and affect expression would develop in the single word period, but with certain cognitive constraints. Namely, affect expressed with words would have positive more often than negative emotional tone, and the words that children said with emotional expression would be among their most frequent and earlier learned words.

**METHODS**

Our subjects were 12 infants, 6 girls and 6 boys, of different ethnic and economic backgrounds, from homes in the New York metropolitan area. All were first-born and their mothers were not employed outside the home at the time the study began. Each infant and mother visited our laboratory playroom once each month from 8 or 9 months to about 30 months of age. Each session lasted one hour. The children were also visited at home every month until they were 15 months old and then every 3 months thereafter. In addition, the mothers kept diaries of the words their children both said and understood at home, in the intervals between monthly playroom observations, and the diaries were reviewed during the home visits. The research reported here was based on data from the video recorded playroom sessions, with the diaries used to determine the children’s earliest words before they said words in the playroom.

The observation playroom was furnished with a child-size table and chairs, a 3-foot plastic slide with a crawl-through tunnel between the steps and the incline, and a changing table. A set of toys was on the floor when the mother and infant entered the room. One of two investigators brought in additional groups of toys at 8-minute intervals according to a schedule, and a snack (cookies with juice for the baby and coffee or tea for the mother) after the first half hour. The same groups of toys were brought into the playroom, in the same sequence, for all the children, and for each child over time. The toys were selected to balance possible girl or boy interest (e.g. doll, truck) and manipulative or enactment play (e.g. nesting blocks, miniature farm animals). We were aware of the possibility that the results of this study might have been influenced by the situation in which the data were collected. For example, the children might have expressed more negative affect if we had seen them in another context, such as bedtime at home, rather than the familiar playroom context in which they and their mothers played with groups of toys and shared a snack. However, the playroom and scheduled presentation of the toys allowed us to control the situation so that it was similar for all the children and for each child over time.

The observations were video-recorded (SLO-383 Sony half-inch stereo Beta). The camera was mounted on a 3-foot high movable tripod in the playroom, and the second investigator maneuvered the camera so that the infant was in view all the time. Each infant and mother interacted with the same pair of investigators throughout the study, both in the playroom and the home visits, and the investigators and infant-mother pairs were matched for ethnicity. The infants and mothers were visited at home before data collection began, and then saw the investigators twice a month; they gave every indication that they were relaxed and comfortable in these playroom sessions.

---

2These were the same subjects as in previous studies of the emergence of language in relation to developments in affect (Bloom et al., 1988; Bloom & Capatides, 1987a) and developments in object play (Lifter & Bloom, 1989).
At the time of recording, a SMPTE time-code generator (FOR-A 3500) imposed an audio signal for each frame of the tape on the second sound track. This allowed the data to be manipulated to within one-thirtieth of a second (each second of video tape containing 30 frames) at the time of data processing. The videotape deck was interfaced at playback with a SMPTE time-code reader and an Apple II+ computer for data coding and transcription. Coders were naive both to the hypotheses to be tested and the units of analysis in the study, and different persons were responsible for coding affect or transcribing speech. The time signal allowed the computer to merge the separate coding passes. This produced an integrated transcription that preserved the temporal relation between speech and affect expression as they occurred in real time. The processed data were transferred to an IBM-XT computer for analyses.

Children vary widely in onset and rate of language development, as is well known. For this reason, we equated the children for language achievement rather than chronological age in order to study the integration of affect expression and developing language. Two language achievements in the single-word period were identified for each child. These were First Words (FW) and a Vocabulary Spurt (VS), both of which have been well-documented in the literature (e.g. Bloom, 1973; Corrigan, 1978; Fischer, Pipp, & Bullock, 1984; Gopnik & Meltzoff, 1987; Nelson, 1973; Stern & Stern, 1907). FW (mean age = 13 months 18 days) was the first observation in the playroom during which a child said one conventional word at least two times. The total number of words used by the 12 children in the 1-hour observation at FW was 63 types ($M = 6$) and 189 tokens ($M = 17$). VS (mean age = 19 months 18 days) was a sharp increase in the number of different words relative to the cumulative slope in the increase of vocabulary from month to month. This was operationalized as the observation in which at least 12 new words were said (an average of 3 per week since the preceding observation) after the child had already acquired at least 20 different words. The total number of words used by the 12 children in the 1-hour observation at VS was 479 types ($M = 40$) and 2,224 tokens ($M = 188$). The data for this study consisted of the speech and affect expression in the first half hour of the two observations in which these developments in language occurred: First Words (FW) and a Vocabulary Spurt (VS).

**Coding Affect Expression**

The emotional signal carries both categorical and gradient information (Stern, Barnett, & Spieker, 1983). Categorical information is the particular emotion (such as anger, joy, or sadness) that is expressed. We are unaware of any explicit theoretical or empirical claims for the relation of categories of emotion to the emergence of language. For this reason, a coding scheme was devised to capture the gradient properties of affective expression, valence (hedonic tone) and intensity, and categories of emotions were not identified or labelled.3

Every change in expressed affect in the stream of an infant’s activities was identified and entered into the computer with the time of onset. Affect coding was continuous, so that the onset time of any change in affect expression was also the offset time of the previous affect expression. This yielded a continuous record of (a) affect expressions, including neutral expression, and (b) the duration of affect expressions from one change in expression to another. An affect expression was any change in affective vocalizations (whining, laughing, and the like), or observable change in the infant’s facial expression, body tension, or posture. Affective vocalisation and facial expression were each considered sufficient by themselves for coding decisions, but body tension or posture required behavior cues from one other channel.

These affect expressions were coded for their *valence*, whether neutral, negative, positive, mixed, or equivocal tone. A neutral expression was defined by the face being in a resting or baseline position as described by Ekman and Friesen (1975), and without body tension or affective vocalization. Mixed affect expressions included elements of both positive and negative valence; equivocal expressions were neither positive, negative, nor neutral, as happened with surprise or excitement. Non-neutral affect expressions were also coded for *intensity* with three levels of intensity indicating the fullness of a display. Thus, the coding scheme for describing the quality of

---

3Neutral affect expression, which predicted early word learning in a previous study (Bloom & Capatides, 1987a), included expression of the emotions category of “interest” (C. Malatesta, personal communication, February 19, 1987). Interest, as an emotion, has been considered important for processes of attention and cognition (Izard, 1986; Piaget, 1954/1981) and for this reason may be a fruitful topic for future research.
expressed affect included three levels of intensity: 1, 2, 3, and five qualities of valence: neutral, negative, positive, mixed, and equivocal. Affective vocalizations such as cries and whines were entered into the coding descriptively.\textsuperscript{4} Ambiguous episodes (the occurrence of a momentary vocalization or facial movement that could not be assigned to one of the above categories) were also coded with time of onset. In addition, the child was sometimes moving away from the camera, or the face was not visible for affect coding with no cues from body tension or affective vocalization. These intervals of “backturn” were also coded for onset so that coding affect expression in the stream of the infant’s activity was not interrupted. Ambiguous affect expression and episodes of backturn were not included in the data analyses. The average amount of coded time, excluding backturn and ambiguous cases, was 24.5 minutes at FW and 24.03 minutes at VS.

\textit{Reliability}

Three coders initially worked in rotating pairs during training, to increase the accuracy of coding and to ensure initial confidence in their judgements. Training continued until the paired coders (the three coders each rotated with one another) achieved a predetermined level of at least 85\% agreement with a segment of data that had been coded by the investigators responsible for training and considered as a standard. Segments were selected randomly from different children for this post-training test of reliability and each segment lasted from 2 to 5.5 minutes. Percentage agreement was computed separately for the categories of valence (positive, negative, neutral, mixed, and equivocal) and intensity (one, two, and three degrees) of nonneutral valence. The level of reliability for the coder pairs, after training and before the actual coding for the study was begun, was high: for valence, Pair 1 = 92\%, Pair 2 = 90\%, Pair 3 = 100\%; and for intensity, Pair 1 = 89\%, Pair 2 = 89\%, and Pair 3 = 85\%.

After continuing to code in pairs for several weeks, reliability was assessed for the individual coders working independently in the following way. Each individually coded a 3.5-minute segment that included at least 25 coding entries. The percentage agreement between members of a pair each coding independently was high: Pair 1 = 100\%; Pair 2 = 94\% and Pair 3 = 94\% for both valence and intensity. At that point the coding was begun for the data used in this study with the coders working independently. The Kappa coefficients of agreement (Cohen, 1960) for valence and intensity ranged from 0.67 to 0.85, for inter-individual reliability and from 0.72 to 1.0 for intra-individual reliability. Cohen’s Kappa corrects for chance agreement and is a more conservative measure than the per cent agreement. These Kappa scores are comparable to those reported in the literature for similar measures and considered evidence of high reliability (e.g. Adamson & Bakeman, 1985). The mean discrepancy in recording affect onset time was an average of 16 video frames, or approximately 0.5 second, for the three coder pairs. (As noted above, the offset of an affect expression was the onset of the next expression.)

\textit{Speech Transcription}

The initial speech transcription was a paper and pencil one. A second transcriber reviewed this transcription in the context of entering it into the computer. Disagreements between the two were resolved by having both review the video recording and transcription together. Lexical items were entered orthographically; nonlexical vocalisations were entered descriptively. All vocalizations were entered with the times of onset and offset.

Accuracy in determining onset and offset times, after training, was high. The mean discrepancy between pairs of independent coders was 2 video frames (1/15 second) for speech onset time, and 5 video frames (1/6 second) for speech offset. (The somewhat lesser accuracy in finding onset time of an affect expression was due to the fact that several kinds of cues were used to code affect, e.g. facial expression and body tension).

\textit{Data Analyses}

Two different measures were used to determine the overlap of speech (discrete events) and affect expression (coded as a continuous variable). The first was a \textit{frequency} measure: The number of times a word was said with

\textsuperscript{4}Photographs of examples of these categories of coded affect expressions are presented in Bloom et al. (1988). The coding scheme, with description of the coding categories and decision rules, is available from the authors. See Schlosberg (1954) and Young (1959), for accounts of the "pleasantness-unpleasantness" and intensity dimensions of emotion; Stern et al. (1983), for discussion of gradient and categorical information in the emotional signal; and Adamson & Bakeman (1982); Ricciuti & Poresky (1972); and Stechler & Carpenter (1967), for other studies that used gradient information in the study of affect expression.
affect expression, coded according to each of the categories of valence (positive, negative, neutral, mixed, and equivocal) and intensity (1, 2, 3). Every speech event (a single word most often) was located in the computerized transcript. The affect expressed at the same time that a word was spoken was identified as follows. With each occurrence of a word, the computer searched backward from the speech onset time until the first onset time of an affect expression was encountered, and forward to the end of a 10-frame window (1/3 second) after speech offset. Within this interval between onset of affect expression and 10 frames after speech offset, affect expression coded with any nonneutral valence took precedence over neutral valence, and higher levels of intensity took precedence over lower levels, for assigning the affect value at the time of speech. For example, if a neutral expression was encountered before speech onset but then an instance of +1 affect expression was encountered either (a) during speech (between speech onset and speech offset), or (b) within the 10-frame window after speech offset, that word was considered to have occurred with +1 affect (emotional expression with positive valence and a low intensity level). The use of the 10-frame window after speech offset was the control for the margin of error in coding times of speech offset and affect onset.

The second analysis was a measure of the time spent in overlap or the number of video frames with overlapping word and affect expression at FW and VS. The number of frames (with 30 video frames per second) was counted for time spent saying words while expressing neutral affect and expressing non-neutral affect.

The temporal relation between saying words and expressing emotion in real time was determined through a form of lag sequential analysis (Bakeman, 1978; Sackett, 1974, 1979). The mean duration of speech (primarily single words but occasionally phrases) was 0.88 seconds (26.3 frames) at FW and 0.85 seconds (25.6 frames) at VS. With the occurrence of a word considered as the anchor lag, 15 1-second lags before and 15 1-second lags after each word were scanned for the occurrence of emotionally toned affect expression. The baseline rate of emotional expression was determined by dividing the amount of time each child spent in emotional expression overall (Bloom et al., 1988) by the total amount of coded time. This was then considered the expected frequency of expression per 1-second lag.

To determine whether emotion was expressed during a lag, the onset and offset of every emotional expression was examined and any overlap with a lag was noted. The lag was considered to have included emotional expression if any overlap occurred. A running count was kept for each instance of emotional expression that occurred in each of the 15 1-second lags before and after a child’s words. The running count was then divided by the number of words to determine the number of times that emotion was expressed in the lag relative to the opportunity for expression. This number, the proportion of lags with emotional expression, was compared to the baseline rate of affect expression overall. The baseline rate of emotional expression was then subtracted from each lag’s proportion of affect to obtain a difference score from baseline. This yielded an emotion expression profile for each child which consisted of the percentages of the difference from baseline in each lag. The differences from baseline were then divided by the standard deviation of the differences for each child (Z-scores) and the means for the Z-scores were used to generate a composite profile for the group of children. This profile was then plotted as the mean difference in standard deviation units from the baseline expression of emotion in each lag.

The remaining analyses concerned those factors that might be expected to contribute to the integration of emotionally toned affect expression and speech at VS. First, to determine the possible differential effects of positive and negative valence, the analysis of overlapping time (number of frames) was performed for positive and negative expression separately. Second, analyses were performed to determine the possible differential effect from the words themselves that the children were learning. The two factors that were examined were (a) earlier word learning, whether words at VS had occurred at FW, had occurred in the interval between FW and VS, or were reported in the mothers’ diaries, and (b) word frequency.

RESULTS

The results consist of (a) the frequency of overlapping word and affect expressions and the time spent in overlapping word and affect expression at the two language achievements, FW and VS; (b) the temporal, lag sequential relation between word and affect expression at FW and VS; (c) the relative time spent in saying words
with positive and negative affect expression; and (d) the effects of early learning and word frequency on emotional expression with speech and the words the children said with overlapping emotional valence.

**Overlapping Word and Affect Expression**

The frequency of overlapping word and affect expression at FW and VS is presented in Fig. 1, as the percent of all words spoken by all the children with negative, neutral, and positive valence, at different levels of intensity. The children said most of their words when they were expressing neutral affect. When they did speak and express emotion at the same time, that emotion was more likely to be level one (-1 or +1) intensity. At VS, more words were said with level 1 intensity and fewer words with levels 2 and 3 than expected given the frequency of emotional expression at the different intensity levels, chi-square = 13.26, \( P < 0.001 \). The same comparison was not significant at FW, \( P > 0.50 \). Thus, the children were not inclined to be speaking at VS when they were expressing emotion with heightened intensity (-2,3, or +2,3).

![Figure 1. Word and Affect Expression Overlap](image)

However, while the children said most of their words with neutral valence, they were also expressing neutral affect most of the time (84% of the time, on average) (Bloom et al., 1988). Thus, one could argue that speech with neutral valence was predicted by the amount of time spent in neutral affect. The second analysis, then, concerned overlapping speech time and emotion time. The number of video frames (with 30 frames per second) was counted for time spent speaking and expressing affect.

**Table 1. Total Time (No. of Frames) in Expressing Affect and Speaking at First Words (FW) and Vocabulary Spurt (VS), \( N = 12 \) children**

<table>
<thead>
<tr>
<th></th>
<th>Speaking</th>
<th>Not Speaking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Affect</td>
<td>Neutral</td>
</tr>
<tr>
<td>FW</td>
<td>3,263</td>
<td>608</td>
</tr>
<tr>
<td>VS</td>
<td>12,850</td>
<td>5,133*</td>
</tr>
</tbody>
</table>

*Chi-square (1 df) = 1.78, \( P = 0.182 \) ns.

*Chi-square (1 df) = 1578.48, \( P < 0.001 \)

*Chi-square (1 df) = 271.63, \( P < 0.001 \).

The amount of time (number of coded frames) spent speaking, not speaking, expressing neutral affect, and expressing emotionally toned affect is presented in Table 1 for FW and VS. At FW, the amount of speech time/emotion time overlap was independent of the amount of emotion time at FW, chi-square = 1.78, \( P = 0.182 \). In contrast, at VS, speech was expressed with emotion more than expected, given the time spent in expressing emotion overall, chi-square = 1578.48, \( P < 0.001 \). Comparing FW and VS, speech time/emotion time overlap increased between the two language achievements, chi-square = 271.63, \( P < 0.001 \), confirming the Integration hypothesis. These results mean that during the moments of speech, the two systems of expression, speech and
emotion, were independent when these infants began to say words. However, by the time they reached the vocabulary spurt, they had integrated the two systems of expression.

**Temporal Relation Between Saying Words and Expressing Emotion**

The results of the lag sequential analyses at FW and VS are presented in Figs. 2 and 3 respectively. The horizontal line in each figure represents the baseline rate or expected frequency of emotional expression per 1-second lag, based upon the amount of time the children spent in emotional expression overall at FW and at VS. The observed frequency of emotional expression is presented as the difference in mean standard deviation units from the baseline expression of emotion in each of the 15 pre- and 15 post-word 1-second lags, plotted in relation to the word lag. Each graph represents the collapsed data from the 12 children, with the resulting curves smoothed by a least squares best fit of the data.

![Figure 2](image1.png)  
*Figure 2. First Words: Profile of emotional expression and saying words (difference in mean s.d. units from the baseline rate of emotional expression in 1-second lags).*

![Figure 3](image2.png)  
*Figure 3. Vocabulary Spurt: Profile of emotional expression and saying words (difference in mean s.d. units from the baseline rate of emotional expression in 1-second lags).*

If the two sorts of expression, emotion and speech, were unrelated to one another, then we would expect an essentially random interaction between expressing affect and saying words. The result would be a lagged function
similar to baseline. However, if the content of a child’s words and the feelings underlying the child’s emotional expressions were about the same thing, as we might expect, then we should find words and emotionally toned affect expression tending to occur together, which is what we found. The expression of emotionally toned affect clustered around the words, but the resulting function had the shape of a sine wave. The critical features of the function were a dip in the frequency of affect expression at about 5 seconds before the word lag, with a steep rise through the word lag and a peak just after the word.

When the lagged functions are compared at FW and VS, the shape of the curves is essentially the same with two differences. The pre-word dip below baseline is steeper at FW, and the post-word peak above baseline is sharper at VS. The occurrence of emotional expression was below baseline in all 15 pre-word lags at FW, but in only 8 of the pre-word lags at VS. The means of the standard deviation units was used as a heuristic to statistically test the differences in relative area at FW and VS. The area beneath the curve below baseline before the word lag was significantly greater at FW ($M = 0.681$) than at VS ($M = 0.177$) ($t (28) = 5.286$, $P < 0.001$). Looking at the 5 lags in the center of each graph, i.e. the word lag with the 2 pre-word and the 2 post-word lags, the area beneath the curve above baseline is significantly greater at VS ($M = 0.889$) than at FW ($M = 0.156$) ($t (8) = -2.306$, $P < 0.05$). These differences are consistent with the results of the analysis of the overlap of time spent in speech and emotional expression which revealed the development that occurred between FW and VS in the integration of affect and word expression. The remaining analyses considered the factors at VS that contributed to this development: the effects of (a) affect valence and (b) the words the children were learning.

**The Valence of Words Said with Emotional Affect**

The words that these infants said with emotional expression at VS were said with positive valence primarily. Time (number of coded frames) spent speaking, not speaking, expressing positive affect, and expressing negative affect at VS is presented in Table 2. Time spent in speech with positive affect expression was greater than was predicted by the time spent in positive affect expression overall, and time spent in speech with negative affect expression was less than was predicted by the time spent in negative affect expression overall, chi-square = 81.06, $p < 0.001$. Thus, the Facilitating Affect hypothesis was confirmed; if these infants expressed emotion and speech together, the emotion they expressed was more likely to be positive than negative.

<table>
<thead>
<tr>
<th>Affect</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speaking</td>
<td>4,277</td>
<td>545</td>
</tr>
<tr>
<td>Not speaking</td>
<td>70,347</td>
<td>13,579*</td>
</tr>
</tbody>
</table>

*Chi-square (1 df) = 81.06, $P < 0.001$

**The Words said with Emotion Expression**

What were the words that these infants said when they were also expressing emotion? The Facilitating Words hypothesis was based on the assumption that the words the children said at the same time they expressed emotionally toned affect would be the words they knew best. To test the first hypothesis, that these words would be their most frequent words, each child’s words were rank ordered according to their overall frequency (number of tokens) at VS. This rank order of overall frequency was compared with the rank order of frequency of words occurring with emotionally toned affect. The rank orders were counted as matching for an individual child when at least two of the three most frequent words overall were included among the four most frequent words said with emotional expression. The second hypothesis was that the words said with emotionally toned affect would have been among the children’s earliest learned words. For this analysis, the words reported in the mothers’ diaries, the children’s words at FW, and words that were said in the interval between FW and VS were examined for the
occurrence of those words said with emotionally toned affect at VS. The effects of frequency and previous occurrence for the individual children are presented in Table 3.

Table 3. Frequency and Previous Occurrence of Words with Emotionally Toned Affect at VS

<table>
<thead>
<tr>
<th>FREQUENCY</th>
<th>PREVIOUS OCCURRENCE with the Interval in months (mos.) between FW and VS in parentheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matching rank orders of word frequency (tokens) overall and emotionally toned words:</td>
<td>Proportion of emotionally toned words (types) at VS which also occurred in previous lexicons:</td>
</tr>
<tr>
<td>8/12 children: Alvin, Charlie, Cory, Diana, Greta, Harry, Reggie, Shirley</td>
<td>More than 75%, 3/12 children:</td>
</tr>
<tr>
<td>Nearly matching, 1 child: Robert</td>
<td>Cory (6 mos.), 82%, Clark (3 mos.), 80%, Robert (6 mos.)</td>
</tr>
<tr>
<td>Not matching, 3/12 children: Clark, Jessica, Vivian</td>
<td>50-75%, 5/12 children:</td>
</tr>
<tr>
<td></td>
<td>64%, Jessica (11 mos), 60%, Charlie; 57%, Shirley (3 mos.), 50%, Greta (7 mos.), 50%, Harry (5 mos.)</td>
</tr>
<tr>
<td></td>
<td>Less than 50%, 4/12 children:</td>
</tr>
<tr>
<td></td>
<td>40%, Diana (5 mos.); 38%, Reggie (9 mos.); 29%, Alvin (9 mos.); 27%, Vivian (3 mos.)</td>
</tr>
</tbody>
</table>

For the 12 children, 8 of the frequency rank orders matched; 1 (Robert) nearly matched (1 word, “Mommy”/“Mama” was ranked highest in both orders); and 3 rank orders did not match. For 8 of the children, 50% or more of the different words they said with emotionally toned affect at VS had occurred in their previous lexicons. Either frequency, previous occurrence, or both predicted which words were said together with emotional expression at VS for all the children except one, Vivian. These results mean that the words these children said with emotional expression tended to be the words they presumably knew well: Words that they said most frequently and/or words they had learned to say at an earlier time.

Included as words in the speech transcription were those words that have traditionally been considered most likely to carry emotional tone, e.g. the interjections such as oh and uoh (Sapir, 1921). However, these interjections were not more likely than other words to occur with emotional expression. A total of 69 different words were said with emotional valence at VS by the group of 12 children. None of these words were names for emotions. Most of the words (47, or 68%) were said by only one child and often only once (22 of the 47 words, or 47%). A small group of words were said with emotion by at least three of the children and these are listed in Table 4. Several of the words in Table 4 are among those that one might have expected would have inherent emotional connotation in addition to their referential meaning for young children. These were (a) the relational words frequently reported in infant vocabularies (no and more); (b) words typically associated with well-known routines (hi, and whee when rolling the ball, using the slide, and the like); and (c) person words (mama and its variants, and baby). Both mama and baby were also said of the small rubber family figures. Of the words used by four or more children, only the highly routinized words hi and whee were said with emotionally toned affect primarily (90% and 79% of the total speech frames respectively).

Finally, to test the conclusion that words said with emotional expression tended to be among the children’s earliest learned and most frequent words, lag sequential analyses were performed with these words excluded from the word lists at FW and VS. For this post hoc analysis, the deviation of each lag from baseline was divided by the maximum deviation in the 31 1-second lags (before, during, and after the words) and the resulting profiles were averaged across children. At FW, the words excluded were (a) words reported earlier in the mothers’ diaries (before the children said their first words in the playroom), and (b) the single most frequent word in each child’s FW lexicon. The observed frequency of emotional expression at FW was below baseline in all but 2 of the 31 lags and, most importantly, neither of the 2 occurred before the word. At VS, words were excluded for the individual
children (a) if they had occurred at FW or in mothers’ diaries, and/or (b) their frequency at VS was 2 standard deviations or more above the mean for the frequency of words overall. At VS, when the most frequent and earliest learned words were excluded, the profiles for 10 of the 12 children had the same shape as the original VS profile in Fig. 3, but differed in amplitude for both the pre-word dip, which was steeper, and the post-word peak, which was not as high. Thus, the expression of emotion was particularly inhibited when these children said words that were not among their most frequent and earliest learned words and, presumably, words that they knew less well.

**Table 4. The Words Said with Emotional Expression by 3 or More Children at VS (N = 12 children)**

<table>
<thead>
<tr>
<th>Word</th>
<th>Number of Children</th>
<th>Word Frequency (Total Tokens)</th>
<th>Proportion of Total Speech Time (Frames)</th>
</tr>
</thead>
<tbody>
<tr>
<td>baby</td>
<td>8</td>
<td>44</td>
<td>0.41</td>
</tr>
<tr>
<td>Mama</td>
<td>6</td>
<td>37</td>
<td>0.43</td>
</tr>
<tr>
<td>hi</td>
<td>5</td>
<td>10</td>
<td>0.90</td>
</tr>
<tr>
<td>whee</td>
<td>5</td>
<td>19</td>
<td>0.79</td>
</tr>
<tr>
<td>more</td>
<td>4</td>
<td>53</td>
<td>0.19</td>
</tr>
<tr>
<td>no</td>
<td>3</td>
<td>16</td>
<td>0.50</td>
</tr>
<tr>
<td>cow</td>
<td>3</td>
<td>45a</td>
<td>0.16</td>
</tr>
<tr>
<td>spoon</td>
<td>3</td>
<td>8</td>
<td>0.60</td>
</tr>
<tr>
<td>three</td>
<td>3</td>
<td>6</td>
<td>0.70</td>
</tr>
<tr>
<td>eye</td>
<td>3</td>
<td>4</td>
<td>0.67</td>
</tr>
</tbody>
</table>

*a*33 of these tokens spoken by one child, Charlie.

**DISCUSSION**

We have asked, in this study, how the two systems of expression available to the 1-year-old child, affect and speech, relate to one another in the period of transition from infancy to language. The results we reported will be discussed in terms of (a) the allocation of cognitive resources for speech and emotional expression; (b) the language of emotion; and (c) the acquisition of language for expression.

**Allocation of Cognitive Resources for Emotional Expression and Language**

The major results of this study were as follows. At First Words, most of the children’s words were said while they were expressing neutral affect, and the time spent in saying words with emotionally toned valence did not differ from what was expected, given the time spent in emotional expression overall. However, by the time they reached a Vocabulary Spurt, toward the end of the period of single-word speech, the children had integrated the two systems of expression. At VS, the time they spent saying words with overlapping emotional expression was greater than was predicted by the amount of time they spent in emotional expression overall.

The integration of the two systems of expression was influenced by several factors. One of these was emotional valence: Words at VS were more likely to be said with positive expression than with negative expression. Another was intensity: Words were more likely to be said with low intensity of emotional expression than with heightened intensity levels. Other factors having to do with the words the children were learning also contributed to the integration of affect expression and speech at VS. The words said at the same time the children expressed emotion tended to be words that the children learned earliest, used most frequently, and/or they were person words (i.e. “Mama” and “baby”) and parts of well known play and other routines. These constraints on saying words together with the expression of emotion can be attributed to the different cognitive requirements that the two systems of expression entail. Positive emotion requires less cognitive work than the negative emotions. Earlier learned, more frequent words are presumably the words that children know best, with the result that they have greatest automaticity and relative ease of expression.
The suggestion has been made that words emerge as the socialized forms of infants’ feeling states, particularly “anxiety” and “affective conflict” (Dore, 1983, p. 168). However, most of the words said by the children we studied were said with neutral affect expression. When they did express emotion and words at the same time, it was with positive affect far more often than negative affect. Others have suggested that words emerge as the socialized forms of infants’ affective vocalizations (e.g. Hilke, 1988). However, the fact that infants can integrate smiles with non-speech vocalizations by the time they are 8 months old, as Hilke has demonstrated, does not predict that words will just naturally follow. Expression through nonlexical vocalizing before speech, and saying words once language begins are not the same. They cannot be equated in terms of their relative cognitive cost to the young language learning child. Words are not effortless. At the minimum, infants need to adapt their own sounds to conform to those in the words they hear; identify sound-meaning correspondences in the language; and recall these correspondences to say a word. Thus, the ability to produce speech with emotional intonation is not innate. While the forms of emotional expression may themselves be innately determined, the ability to integrate these forms with speech is learned.

If the cognitive activity needed for saying words competes with the cognition required for the experience and expression of emotion, one result would be suppression of emotional expression at times of word learning and/or use. At a micro level, the results of the lag analyses showed such a suppression of emotionally toned affect approximately 5 seconds before the occurrence of a word. We are interpreting the pre-word dip below the baseline level of emotional expression (in Figures 2 and 3) as the time in which the mental activity associated with the expression of emotion is, essentially, suspended. This is the time the child uses for the cognitive work entailed in recalling and saying a word. This effect was heightened for words that were relatively infrequent, learned most recently, and so presumably less well known.

The lag analyses in Figs 2 and 3 also show that after the post-word peak in emotional expression, the expression of emotion once again dipped below baseline. Although we have not yet looked systematically at other events surrounding the children’s words, we know that their mothers invariably responded to what they said. We speculate that this dip below baseline level of emotional expression in the post-word lags reflects an inhibition in emotional expression that can be attributed to the children’s listening to their mothers’ responses to what they themselves had just said.

The peak in affect expression comes after having said the word, when the child’s cognitive resources are freed for the cognitive work that the experience and expression of emotion involves. Given that most of the children’s emotional expression was positive, this burst of emotional expression just after a word is reminiscent of the smile of recognition observed with younger infants by McCall (1972) and Zelazo (1972), and smiles following mastery (Sroufe & Waters, 1976) or assimilation after “concentrated attention” (Kagan, Lapidus, & Moore, 1978). Thus, words and emotional expression were closely associated in time, but emotion was more likely to be expressed in the seconds immediately after a word. The cognition required for the experience and expression of emotion could be expected to pre-empt the cognitive activity required for learning words as well. At a more macro level, we have reported elsewhere that more frequent expression of emotion did not enhance early language acquisition for the children in this study (Bloom & Capatides, 1987a).

Those children who expressed emotionally toned affect less often and, correspondingly, spent more time in neutral affect expression, reached achievements in language (first words, a vocabulary spurt, and the use of multi-word speech) earlier than the children who expressed emotion more frequently. A similar result for language acquisition has been reported as well by Epport (1987) and Nachman (1986). A related finding is that higher ratings by mothers of smiling and laughter at 8 and 12 months of age was correlated with lower developmental test scores (the mental development index of the Bayley Scales of Infant Development, which includes several measures relevant to language) in a study by Fagen, Singer, Ohr, & Fleckenstein (1987).

In addition, learning language appeared to have a stabilizing effect on emotionality as a dimension of temperament. The earlier word learners in our study were more stable over time in their profiles of emotionality than were the later word learners (Bloom & Wikstrom, 1987). For example, all the correlations (Pearson r) for time spent in positive affect expression at 9, 13, 17, and 21 months were significant ($P = 0.05$ or less) for the earlier word learners, and the ages 13, 17, and 21 months coincided closely with their mean ages of language
In sum, we propose that the different cognitive requirements for emotional expression and language compete for the young language-learning child’s attentional resources. Competition for the resources required for (a) the experience and expression of emotion, and (b) expressing a word at the same time that articulates what the emotion is about, was reflected in the temporal relation between saying words and expressing emotional affect. While the expression of emotion is generally considered to be automatic and effortless, the children in this study did not, in fact, express emotion just before saying words.

The findings reported here are consistent with accounts in the neuropsychological literature concerning different hemispheric specialization in the brain for language and emotion. Separation of functions in the different hemispheres of the brain results in a competition for attention and activation when tasks require both functions (Kinsbourne, 1975). We have proposed that saying words with emotionally toned valence is such a task that is particularly difficult for the young language-learning child, especially given the relatively immature status of the late-maturing corpus callosum. The question of integrating the results of separate processing by the two hemispheres of the brain was raised by Tucker (1986) who suggested that “lateralized contributions to communication . . . may develop before the child has the capacity to coordinate them” (p. 273). However, we expect that the effects we observed are not limited to the earliest stages of language learning. Older children and even adults may well experience the same competition between the two systems of expression for cognitive resources at times when the experience of emotion is particularly heightened and/or the content of speech is particularly abstract, complex, or obscure.

The Language of Emotion

The children in this study learned words that could express what their feelings were about in this period of time before they learned the actual names for the emotions themselves. Acquisition of words for naming emotions and feeling states is a relatively late development (Beckwith, 1989), even in cultures such as Samoa where talk between children and their mothers is “intensely emotional” (Ochs, 1986, p. 252). Reports in the literature of the “early” use of emotion words by English-speaking children have not been observational but have relied instead on mothers’ reports of whether such words occur at all, using checklists. The youngest children in these studies have generally not been less than 20 months old (e.g. Bretherton, Fritz, Zahn-Waxler, & Ridgeway, 1986).

Similarly, the mothers of the children whom we studied used emotion words very infrequently when talking in the moments that surrounded their children’s emotional expression. Instead, they talked about the situations that were the causes, consequences, and circumstances of their children’s emotional expressions (Capatides, 1989). When the mothers’ speech before and after each affect expression by the child was examined, the names for the emotions (e.g. happy, mad,) were rare and actually decreased in frequency in the period from 9 to 21 months. The mean number of emotion word tokens for the 12 mothers, in 30-minute observations, was 2.5, at 9 mos; 1.17, at 13 mos; 0.83, at 17 mos, and 0.5 at 21 months. In fact, these frequencies include the most frequent affect word the mothers used, like, which may have marginal status as an emotion term.

Mothers’ conversations about “feeling states” were described by Dunn, Bretherton, & Munn (1987) for children in the same age range as our study. In the list of “feeling-state labels” that the mothers in their study used, “emotional states” was one of three categories (along with “quality of consciousness” and “sensations and physiological states”). Similarly, the number of names of emotions in the category of “emotional states” was extremely small and most of the words the mothers used in this category (e.g. enjoy, mind, temper, laugh, and the like) were not emotion terms in the sense described, for example, by Clore and Ortony (1987). In any event, the mean number of mother utterances referring to any of the three categories of feeling states in their study, pro-rated for 30 minutes for the sake of comparison, was 1.75 and 2.75 at 18 months and 24 months respectively.
In sum, children’s early words in the second year of life do not name their emotions and so they cannot tell us what they are feeling. Rather, they continue to rely on expression of emotion through facial and postural displays of affect while the words they are learning express what their feelings are about. Moreover, the language that 1-year-old children hear instructs them in the sorts of experiences that are associated with different feeling states, more often than it provides labels for the emotions themselves (Capatides, 1989; Dunn, Bretherton, & Munn, 1987).

The Acquisition of Language for Expression

In the present study, the two forms of expression, affect and speech, occurred closely in time. We can assume, then, that the words these children were learning in this period were words for expressing something of the objects and circumstances of their emotional experience. This result supports the proposal that language is learned to express what children have in mind, the contents of attention and feelings, and children learn language for expression rather than as an instrument for getting things done in the world (Bloom & Beckwith, 1988). To be sure, young children’s words influence what other persons do, but that effect is subordinate to the expressive power that the infant achieves through language. The power of expression comes with the ability to take something that is hidden within the individual and make it manifest, to put it in a public place (Taylor, 1979). Contents of feelings, beliefs, and desires are hidden within the individual. Affect and language are two modes of expression that allow us to attribute such internal properties to the expressor. Affect can express how the infant feels about such mental contents. By learning language, the infant comes to express and articulate what these contents are and to attribute such hidden mental contents to others based upon what they understand them to say.

Cognitive theories of emotion stress the importance of goals and plans, and changes in the environment that influence success or interfere with an individual’s goals or plans. If these are the facts that produce emotional experience (e.g. Oatley, 1988) and underlie knowledge about what emotions are (e.g. Stein & Levine, 1989), then learning to talk about such items of experience is at least as important as learning the emotion terms themselves. In fact, to the extent that the infant’s facial expression and postural displays of emotion are available to others, emotion terms would be more redundant than informative. We have already shown in other research that expression having to do with subjective, emotion-eliciting events contributes to how children learn the language of causality (Bloom & Capatides, 1987b; Hood & Bloom, 1979). Children begin to talk about the causal connections entailed in emotional experience even before they learn such linguistic terms for causality as because, so, and why. Similarly, the children in the present study learned the language for expressing the causes and objects of their emotions which had to do, no doubt, with the success or failure of their goals and plans, before they even began to acquire a lexicon of emotion terms.

While the feeling state itself is not named by words, at least part of the contents attributable to the expression of affect in the first year is now made manifest with words in the second year. But, in addition, as children approach the second year of life, all they have learned about objects, persons, and the self inform their beliefs and desires, and these cannot be expressed by affect alone. Other modes of expression are required. Waiting in the wings, so to speak, is language. Language is the pre-eminent mode of expression and is provided by society and culture for making manifest and public that which is internal to the individual. The infant’s cognitive, social, and biological resources are soon bent to the task of its acquisition.

Language does not replace affect expression and children continue to express their feelings through affect as they learn language (Bloom et al., 1988). Indeed, with further development, they will also become aware of new and different feeling states that require increasingly more subtle and controlled forms of expression. These include the more socially and personally complex emotions (for example, jealousy, guilt, and shame). Language can express many aspects of the objects, circumstances, and feelings associated with the more complex emotions, but by no means all their aspects. For this reason, the integration of affect and language will continue to develop for expression in the increasingly wider contexts of the child’s development.
REFERENCES


