Infant and Child Development
Inf. Child Dev. **9**: 1–16 (2000)

Infant Vocal Development during the First 6 Months: Speech Quality and Melodic Complexity

Hui-Chin Hsu[a,*], Alan Fogel[b] and Rebecca B. Cooper[b]
[a] *Department of Child and Family Development, The University of Georgia, Athens, GA, USA*
[b] *Department of Psychology, University of Utah, Salt Lake City, UT, USA*

The early development of infant non-distress vocalizations was investigated in this study. Thirteen infants, from 4 to 24 weeks of age, and their mothers were observed weekly in a face-to-face interaction situation. The speech quality (syllabic versus vocalic) and melodic complexity (simple versus complex) of infant vocalizations were coded independently. Based on speech quality and melodic complexity, four types of infant non-distress vocalizations were categorized: simple and complex syllabic (speech-like) vocalizations as well as simple and complex vocalic (non-speech-like) vocalizations. Results showed that complex syllabic sounds were of longer duration and complex vocalic sounds were less frequent than the other types of vocalizations. Curvilinear developmental trends were found in the rate of simple vocalic sounds and in the mean duration of complex syllabic sounds. Furthermore, before 4 months of age, vocalic sounds were more likely to be associated with simple melodic contours, after which syllabic sounds were more likely to be associated with complex melodic contours. A dynamic systems perspective on the early development of infant vocalization is discussed. Copyright © 2000 John Wiley & Sons, Ltd.

Key words: dynamic systems theory; infant vocalization; non-linear development

Early infant vocalization is viewed as a precursor to interpersonal communicative skills and language development (Bateson, 1975; Stark, 1978). Because of dramatic changes in infant vocalization during the first year, the development of various types of infant vocalizations such as cries (Prescott, 1975), laughter (Nwokah *et al.*, 1994), grunts (McCune *et al.*, 1996), and non-distress vocalizations (e.g. Stark, 1978) have been examined extensively. Numerous stage-like models have also been formulated to identify and describe the emergence of qualitatively different vocalizations and the achievement of developmental

* Correspondence to: Department of Child and Family Development, The University of Georgia, McPhaul Center, Athens, GA 30602–3622, USA.
E-mail: hchin@fcs.uga.edu

Received 20 May 1999
Accepted 23 September 1999

milestones (cf. Vihman, 1996). Empirical studies demonstrate that vocal development in infancy follows an orderly sequence (e.g. Koopmans-van Beinum and van der Stelt, 1986; Stark *et al*., 1993). On closer examination, however, research further shows that earlier forms of vocal behavior do not simply disappear as new ones emerge (Stark *et al*., 1993). This indicates that early infant vocal development may not be a series of stepwise jumps. Rather, it may be a continuous, but non-linear, process.

The principle of self-organization from a dynamic systems perspective may best explain the continuity and non-linearity in early vocal development. The dynamic systems approach argues that vocal production is the outcome of coordinations among articulatory, respiratory, and anatomical elements (Kent, 1984; Thelen, 1991). Owing to the constraints imposed by the configuration of these elements, the vocal system retains its invariance or continuity over time. When changes occur in internal (e.g. maturation of vocal tract) or external (e.g. social environment) conditions, elements of the vocal system reassemble themselves. A new configuration of these elements results in a non-linear shift in the vocal system. This phase shift does not necessarily lead to progressions (e.g. increment in occurrence and/or enhancement in quality) but sometimes regressions (e.g. attenuation in occurrence and/or degeneration in quality) of the system. Despite the theoretical recognition of continuities in early vocal development (Stark, 1978), inherent in the traditional stage-like developmental models of infant vocal production is the underlying assumption of linear progression (versus regression) and stepwise discontinuities (versus continuities). Very little research effort has been made to examine the continuity and non-linearity of vocal development in early infancy. Guided by a dynamic systems perspective of development, this study examined developmental changes in the quantitative parameters (e.g. duration, melodic contours, and fundamental frequency) of infants' non-distress vocalizations with various sound qualities and prosodic patterns.

Speech quality (i.e. segmental sound quality) and melodic complexity (i.e. suprasegmental prosodic features) are two primary characteristics of early non-distress vocalizations. With respect to speech quality, the perceived speech-likeness in infant vocalizations by adult listeners is the main concern. A vocalization, for example, is classified as speech-like when it contains vowel- or consonant-like qualities (Maskarinec *et al*., 1981). A metaphonological model, proposed by Oller (Oller, 1986; Oller and Lynch, 1992), further suggests that the evaluation of vocal quality in early infancy is based on the degree to which vocalizations approximate the infrastructural characteristics of adult speech. As a result, in the first few months of life, two types of infant vocalizations with different speech quality can be identified: (1) a fully-resonant vowel sound with normal phonation, and (2) a quasi-vowel sound characterized by a nasal anti-resonance with low pitch.

Based on the resonance pattern (oral or nasal), the sound location (anterior or posterior area of the mouth), and the perceived effort in sound making (relaxed or forced), Bloom (1988, 1989, 1993) classified infant non-distress vocalizations into syllabic and vocalic sounds. Compared with vocalic sounds, syllabic sounds are vocalizations with more speech-like quality. Using non-distress vocalizations sampled from 3-month-old Canadian and Japanese infants, acoustic analysis demonstrated that syllabic sounds were longer in duration and that vocalic sounds were more nasal (Bloom, 1988; Masataka and Bloom, 1994; Masataka, 1995). A number of experimental studies have shown that young infants exhibit more vocalic than syllabic sounds when the adult partner becomes unresponsive or non-contingent (Bloom *et al*., 1987; Legerstee, 1991). An increase in the production of syllabic sounds was found when interacting with a responsive and contingent adult partner, which suggests that the speech quality in infant non-distress vocalizations varies consistently with the quality of social interaction (Bloom *et al*., 1987; Bloom, 1988; Legerstee, 1991).

The developmental pattern of infant non-distress vocalizations during the first 6 months of life has been examined in a number of longitudinal studies. Several of these studies were designed to document the pre-speech development of normal developing infants in comparison with babies born with abnormalities (e.g. prematurity, Down Syndrome, deafness, etc.). Some of these studies aimed to describe the effects of physical and social contexts on the acoustic features (e.g. duration, melodic contours, and fundamental frequency) of infant vocalizations over time. Developmental change in the duration and the rate of infant vocalizations with different speech qualities, however, has not been systematically examined. Despite longitudinal designs, vocal samples in these studies were collected from various contexts and various methods were used in measuring the length or rate of infant vocalizations. In general, the mean duration of overall infant vocalizations (typically shorter than 400 ms and ranging from 220 to 800 ms) increases over time (Laufer and Horii, 1976; Delack, 1978; Laufer, 1980), whereas the overall rate shows an inverted-U pattern (Tipps *et al*., 1981; Berger and Cunningham, 1983; Legerstee, 1991). Nevertheless, the developmental patterns found in most of these studies were not subjected to rigorous statistical testing. Only one study applied a linear modeling test. The methods used, and the results found, in these studies are summarized in Table 1.

With regard to prosodic features of infants' non-distress vocalizations, melodic contours are of primary interest. Infants can imitate (Kessen *et al*., 1979), discriminate (Trehub *et al*., 1984), and produce (Papousek and Papousek, 1981) a variety of melodic patterns before they can speak. It has been suggested that early vocal development is primarily manifested in melodic contours during the first 6 months (Tonkova-Yampol'skaya, 1973). D'Odorico (1984) reported that 4–9-month-old infants exhibited vocal sounds with different melodic patterns for different communicative functions. Request sounds, for example, were characterized by flat (level) as well as rising pitch contours, call sounds were more likely to show a rising contour pattern, and discomfort sounds tended to have flat or falling pitch contours. Delack and Fowlow (1978) found an association between melodic contours and the contexts in which the vocalization occurred. The rise–fall contour was more likely to occur when infants were interacting with their mothers, whereas the rise and fall–rise contours were more likely to occur when infants were alone. Bloom (1988, 1989, 1993) also reported that speech-like syllabic sounds exhibited by 3-month-old infants showed more variable pitch patterns than non-speech-like vocalic sounds. Nevertheless, it is not known whether there are developmental changes in the mean duration and rate of infants' non-distress vocalizations with different melodic contours. Furthermore, it is not clear whether there are developmental changes in the association between speech quality and melodic complexity of infant non-distress vocalization during the first 6 months of life.

The speech quality and melodic pattern of infant non-distress vocalizations can be assessed by objective and perceptual methods. Objective methods include articulatory and phonatory methods (Koopmans-van Beinum and van

Table 1. Summary of longitudinal studies on the development of infant vocalizations during the first 6 months of life

Research project	Sample size	Interval of data sampling	Contexts of data sampling	Measures of vocal samples	Data analysis of developmental pattern
Berger and Cunningham (1983)	15 (8 Down Syndrome infants)	Weekly; began at a mean age of 5.3 weeks	Sampled 40-s mother–infant face-to-face interaction and mother silent condition; at home	Event recording of non-distress vocalizations	Visual inspection
Delack (1978), Delack and Fowlow (1978)	19	Biweekly; began at 1 month	Alone, with objects, with mother, and with an experimenter; at home	Spectrographical analysis of all spontaneous vocalization	Visual inspection
Kent and Murray (1982)	21	3, 6, and 9 months	1-h interaction with mother and experimenter; at a laboratory	Spectrographical analysis of non-vegetative sounds	Visual inspection
Laufer (1980)	4	Biweekly; began at 1 week	Parent–infant caregiving routine at home; alone, dyadic parent–infant, and triadic mother–father–infant interaction at a laboratory	Spectrographical analysis of proto-syllabic sounds	Visual inspection
Laufer and Horii (1976)	4	Biweekly; began at 1 week	Parent–infant caregiving routines at home; alone, dyadic parent–infant, and triadic mother–father–infant interaction at a laboratory	Spectrographical analysis of non-distress and non-vegetative sounds	Visual inspection
Legerstee (1991)	8	Biweekly; began at 3 weeks	45–60 s interaction with mother, experimenter, and object at a laboratory	Unspecified method in measuring non-distress sounds	Analysis of variance and visual inspection
Maskarinec *et al.* (1981)	5 (2 infants with hearing impairment)	Weekly (3 times); began at 1–11 weeks	10-min recordings at home and at a laboratory (parent and experimenter were silent)	5-s block time sampling of non-speech-like and speech-like sounds	Linear regression modeling
Oller *et al.* (1994)	53 (33 preterm infants)	Bimonthly; began at 4 months	30-min recording at a lab room (parent and experimenter elicited vocalization through social interaction	Infraphonogical categorization (i.e. approximation of mature speech) of non-distress sounds	Analysis of variance
Stark (1978, 1989)	2	Weekly or biweekly; began at 1 and 2.3 months	Mother–infant caregiving routines at home	Spectrographical analysis of cry, discomfort, vegetative, and cooing (comfort) sounds	Visual inspection
Tipps *et al.* (1981)	1	Weekly; began at 3 weeks	5-min interaction with parents and objects, and during motor activities; at home	Used paper-and-pencil measuring speech and non-speech sounds	Visual inspection

der Stelt, 1986), as well as phonetic and acoustic analysis (Nakazima, 1975; Stark, 1980, 1981; Roug *et al*., 1989). These analyses are extremely labor intensive, and specialized training of the coders is usually required. Typically, vocal samples from a few infants are collected for analysis. Auditory perceptual assessments by human observers, on the other hand, provide subjective evaluations of speech parameters. Inferences about speech production processes and attributions of traits or underlying emotional states in the vocal sounds can be made (Scherer, 1982). Furthermore, metaphonological characteristics such as smoothness and timing of transitions (Oller, 1978, 1980; Oller and Lynch, 1992) of a vocalization can only be evaluated by perceptual analysis.

Recent evidence suggests that adult listeners can reliably perceive the speech quality of infant vocal sounds (Oller and Lynch, 1992) and infer infants' communicative messages (Papousek, 1992). Bloom and her colleagues (Bloom, 1993; Beaumont and Bloom, 1993; Bloom *et al*., 1993) further demonstrated that both parents and non-parents were more favorable to infants who produced speech-like syllabic sounds. Such infants were perceived as more attractive, friendly, and sociable than those who exhibited less speech-like vocalic sounds. Infants who produced syllabic sounds were also perceived as more intentional in their efforts to vocalize (Bloom, 1993; Beaumont and Bloom, 1993). Mothers were also verbally more responsive to speech-like syllabic than to non-speech-like vocalic sounds (Masataka and Bloom, 1994). These results suggest that it is the subjective perception, rather than the objective acoustic or phonetic identification of infant vocal signals, that is more meaningful and ecologically valid for an adult partner of an infant during social interaction.

The present study was designed to investigate developmental changes in infant non-distress vocalizations during social interaction during the first 6 months using perceptual analysis. The first goal of the study was to examine developmental changes in quantitative parameters (such as duration and rate) of infants' non-distress vocalizations, varying in speech quality and melodic complexity and produced during mother–infant interaction during the first 6 months. Based on previous empirical work on infant vocal development, we expected that there would be developmental changes in the duration and rate of infants' non-distress vocalizations. Based on previous findings on the duration and rate of infant overall vocalization, it was predicted that, regardless of the speech quality and melodic complexity, a linear increment in the duration of non-distress vocalizations and a non-linear developmental trajectory with the rate of non-distress vocalizations would be found. The second goal of this study was to examine the pattern of association between speech quality and melodic complexity in infants' non-distress vocalizations during the first 6 months of life. Even though research suggests developmental changes in both speech quality and melodic contours, there is no evidence to suggest that the pattern of association between the two would change over time. Therefore, based on Bloom's studies with 3-month-old infants, we predicted that speech-like sounds would be more likely to exhibit complex melodic contours, whereas non-speech-like sounds would be more likely to manifest simple melodic contours over the course of the first 6 months of life.

METHOD

Subjects

Thirteen mother–infant dyads participated in a 2-year longitudinal study on infant communication development. All the infants were fullterm with no major birth complications, were from intact families, and passed a hearing test at 6 months of age. Twelve mother–infant dyads were Caucasian, and one was African-American. Six of the infants were first-born; eight were male and five were female. Nine of the mothers had a Bachelor degree, two had some college education, and two had their high school diploma. Seven of the mothers were employed full-time, two were employed part-time, and four were homemakers. Because of the missing data for two infants during the first month, only 11 subjects were included in the Analysis of Variance (ANOVA) procedure (see below).

Procedure

Infants and their mothers were videotaped in a laboratory room beginning when the infants were between 4 and 9 weeks of age ($M = 5.3$ weeks). Infants and their mothers were videotaped weekly during the first year, and then biweekly during the second year, under four conditions (including lap, floor, high chair, and table conditions; for details see Nwokah *et al*., 1994). The mothers were instructed to play with their infants as they normally would at home. Only the face-to-face interaction data from the lap sessions during the first 6 months were included in the current study. Mothers were seated on a straight-back chair and held their infants in their laps. No toys were provided in this condition. Sessions lasted approximately 5 min, except when the infant became too fussy to continue (nine out of total 213 sessions). The average session duration was 287 s (range = 80–300 s). The average number of sessions collected from each dyed was 16 (range = 9–20).

Three remote-controlled cameras were used to film the play sessions. One camera was focused on the mother's upper body and the side of the infant, and the other two cameras were both focused on the infant's face and body. The outputs from the two cameras which had the best views of the mother and the infant were passed through a special effects generator to produce a split-screen image with a timer (accurate to 0.01 s) superimposed on the screen. A microphone (Shute 575SB), hung from the ceiling about 12 in. above the mother's head, transmitted the audio signals to an amplifier (Shure M267) for recording.

Coding

The primary coder was a professionally trained vocalist and an experienced mother majoring in psychology. The onset and offset times of a non-distress or non-vegetative vocalization were coded from the video. The coder listened to the sound until a non-distress vocalization was made by the infant. The coder then stopped the tape and recorded the onset time. The procedure was repeated several times to obtain the offset time of a vocalization and to recheck the recorded times. In line with previous studies (e.g. Stark, 1978; Oller *et al*., 1994), vegetative sounds (such as wheezes, sneezes, coughs, hiccups, and clicking sounds), effort and grunting sounds, negative vocalizations (such as whimpering, fusses, and cries), and laughs were excluded.[1] No babbling (i.e. sequences of consonant-like and vowel-like sound; MacNeilage and Davis, 1992) was observed in the collected samples. A vocalization was identified as a discrete sound occurring within one respiration that is intuitively defined, consistent with the procedure outlined in previous research (Maskarinec *et al*., 1981; Oller and Lynch, 1992). If the sound was segmented by an interruption of phonation (i.e. a perceivable silence), then two separate sounds were recorded. A reliability coder independently coded 16% of the total sessions for the duration and the rate of infant vocalizations. The mean differences of the duration of vocalizations (syllabic = 0.05 s and vocalic = 0.07 s) between the two coders were computed. The intraclass correlations were also computed for the duration (syllabic = 0.93 and vocalic = 0.90) and rate per minute (syllabic = 0.87 and vocalic = 0.76) of infant vocalizations.

Following the determination of the onset and offset times of vocalizations, the primary coder replayed the video segments and judged the quality and the melodic complexity of the vocalization. A total of 1692 infant non-distress vocalizations were coded. In order to achieve more reliable estimates, the data were collapsed into five monthly periods: (1) 2nd month: 4–8 weeks; (2) 3rd month: 9–12 weeks; (3) 4th month: 13–16 weeks; (4) 5th month: 17–20 weeks; and (5) 6th month: 21–24 weeks. An independent coder for speech quality coded a total of 16% of the sessions. The two categories of speech quality and the two categories of melodic complexity of these vocalizations were decided separately. The coding procedures and reliability testing are reported in the following sections.

Speech quality

The quality of infant vocalization was evaluated and determined by the perceived speech-likeness in the sound. All the vocalizations were categorized as either syllabic or vocalic according to Bloom's (1988, 1989, 1993) definitions and classification. *Syllabic sounds* were defined as those uttered in the anterior area of the mouth containing greater oral resonance and were perceived as more relaxed and speech-like. *Vocalic sounds* were defined as those produced in the posterior area of the mouth containing greater nasal resonance and lacking projection, and were perceived as more forced and less speech-like. A third coder coded 16% of the session for reliability checks on the coding of speech quality. The reliabilities were calculated at monthly period: Kappas were 0.99, 0.73, 0.82, 0.90, 0.69, and percentages of agreement were 99%, 78%, 86%, 90%, and 88%, respectively.

Melodic contours

Like previous studies on the meaning of vocalizations in naturalistic social communication (e.g. Marcos, 1987), the melodic contours of infant vocalizations were coded perceptually in this study. Melodic complexity was judged based on a classification system used by Papousek and Papousek (1989): flat, rising, falling, bi-level U-shape, inverted-U shape, and sinusoidal. Because of low frequencies with some of the melodic patterns, the categories were further collapsed in order to simplify the data analysis. The first three melodic patterns (flat, rising, falling) were collapsed as *simple* patterns, and the last four (U-shape, inverted-U shape, bi-level, and sinusoidal) were collapsed as *complex* patterns. Fourteen per cent of the sessions were coded by a fourth coder for reliability checks on the coding of melodic contour. The reliabilities were also computed at monthly period: Kappas were 0.67, 0.62, 0.58, 0.65, and 0.61, and percentages of agreement were 79%, 71%, 69%, 74%, and 75%, respectively.

RESULTS

Developmental changes in quantitative parameters

To examine developmental changes in the quantitative parameters of infant non-distress vocalizations, the mean duration (s) and the rate per minute of infant vocalization were derived at monthly level and used as dependent variables in the two separate 2 (speech quality: syllabic versus vocalic) $\times$ 2 (melodic complexity: simple versus complex) $\times$ 5 (monthly period) repeated-measures ANOVAs. Because the observations for two infants did not begin until they were 9 weeks old, as a result of casewise deletion, the sample size for the repeated-measures ANOVA was 11 instead of 13. Owing to the small sample size, the Greenhouse–Geisser adjustment for the significance test was employed.

Mean duration

Results from the repeated-measures ANOVA showed that there were significant main effects of speech quality and melodic complexity. The two-way interaction effects of speech quality $\times$ melodic complexity and speech quality $\times$ monthly period were also significant. The mean duration of syllabic sounds were longer than those of vocalic sounds ($F(1,10) = 13.21$, $p < 0.005$), whereas vocalizations with complex melodic contours were longer than those with simple ones ($F(1,10) = 14.92$, $p < 0.004$). *Post hoc* comparisons of the means showed that complex syllabic sounds were significantly longer than the other three types of vocalization at 3, 4, and 6 months ($F(1,10) = 6.66–18.69$, $p < 0.002–0.02$) (see Table 2), but not at 2 and 5 months.

To further detect developmental trends, polynomial analyses were performed. The results demonstrated that the mean duration of complex syllabic sounds

Table 2. Means and S.D.s (in parentheses) of infant non-distress vocalization by monthly periods: mean duration and rate per minute

Variables	Vocalization types	Monthly periods				
		2	3	4	5	6
Mean duration (s)	Simple syllabic	0.26 (0.23)	0.41 (0.27)	0.48 (0.30)	0.34 (0.18)	0.34 (0.21)
	Complex syllabic	0.31 (0.51)	0.70 (0.60)	0.91 (0.52)	0.55 (0.43)	0.72 (0.37)
	Simple vocalic	0.40 (0.23)	0.41 (0.20)	0.38 (0.21)	0.35 (0.24)	0.33 (0.26)
	Complex vocalic	0.41 (0.33)	0.38 (0.34)	0.53 (0.47)	0.40 (0.45)	0.41 (0.81)
Rate per minute	Simple syllabic	0.37 (0.56)	0.51 (0.59)	0.56 (0.55)	0.63 (0.71)	0.43 (0.54)
	Complex syllabic	0.45 (0.90)	0.57 (0.50)	0.63 (0.69)	0.60 (1.06)	0.49 (0.47)
	Simple vocalic	0.42 (0.43)	0.60 (0.45)	0.68 (0.65)	0.33 (0.22)	0.31 (0.25)
	Complex vocalic	0.21 (0.34)	0.23 (0.27)	0.25 (0.22)	0.21 (0.30)	0.09 (0.10)

exhibited a significant non-linear cubic ($F(1,10) = 7.31$, $p < 0.03$) and quartic trend ($F(1,10) = 5.33$, $p < 0.05$) (see Figure 1(a)). The mean duration reached its peak at the 4th month and declined at the 5th month. After the decline, there was a slight increase at the 6th month. No significant developmental changes in the mean duration were found for the other three types of non-distress infant vocalization.

Rate per minute

Results from the repeated-measures ANOVA revealed that there were significant main effects of speech quality and melodic complexity. The two-way interaction effect of speech quality × melodic complexity was also significant. Infants produced significantly more syllabic than vocalic sounds ($F(1,10) = 7.17$, $p < 0.03$) and significantly more vocalizations with simple melodic contours compared with those with complex ones ($F(1,10) = 6.89$, $p < 0.03$). *Post hoc* comparisons showed that complex vocalic sounds were significantly less frequent than the other three types of vocalizations at 2, 3, 4, 5, and 6 months of age ($F(1,10) = 5.16–21.6$, $p < 0.001–0.05$) (see Table 2).

Polynomial analyses were also performed to detect developmental trends. The results showed that the rate of simple vocalic sound exhibited a significant non-linear quadratic trend ($F(1,10) = 5.70$, $p < 0.04$) (see Figure 1(b)). The rate of infant simple vocalic sounds reached its peak at the 4th month, and then declined. No significant developmental changes were found for the rate of the other three types of non-distress infant vocalization.

In summary, infant vocalizations with different types of speech quality and melodic complexity were significantly different in their quantitative parameters, including mean duration and rate per minute. Complex syllabic sounds were the longest in duration, whereas complex vocalic sounds were the least frequent. Non-linear developmental changes involving both progression and regression were also found in the mean duration of complex syllabic sounds and the rate of simple vocalic sounds.

Developmental changes in association between speech quality and melodic complexity

To examine developmental changes in the association between speech quality and melodic complexity across time, a 2 (speech quality: syllabic versus vocalic) × 2 (melodic complexity: simple versus complex) × 5 (monthly period) multidimensional contingency table was formulated based on the rate of infant vocalization collapsed across all the infants. A log-linear analysis was then performed. The results revealed that the best fitting model included the three main factors and two two-way interaction terms: speech quality × melodic complexity and speech quality × monthly period ($G^2(8) = 9.63$, $p = 2.92$).

To identify the specific patterns in the cross-tabulation, a follow-up Configural Frequency Analysis (CFA; Lienert and Krauth, 1975; von Eye, 1990) was performed. CFA is a method for the analysis of the individual cells formed by the cross-tabulation in a multidimensional contingency table. CFA compares observed with expected frequencies in order to test the strength of the association. The results of the analysis revealed that, before 4 months, there was an association of vocalic sounds with simple melodic contours and a dissociation of syllabic sounds from simple melodic contours. However, after 4 months of age, an association was found between syllabic sounds and complex melodic contours, whereas the dissociation pattern was found between vocalic sounds and

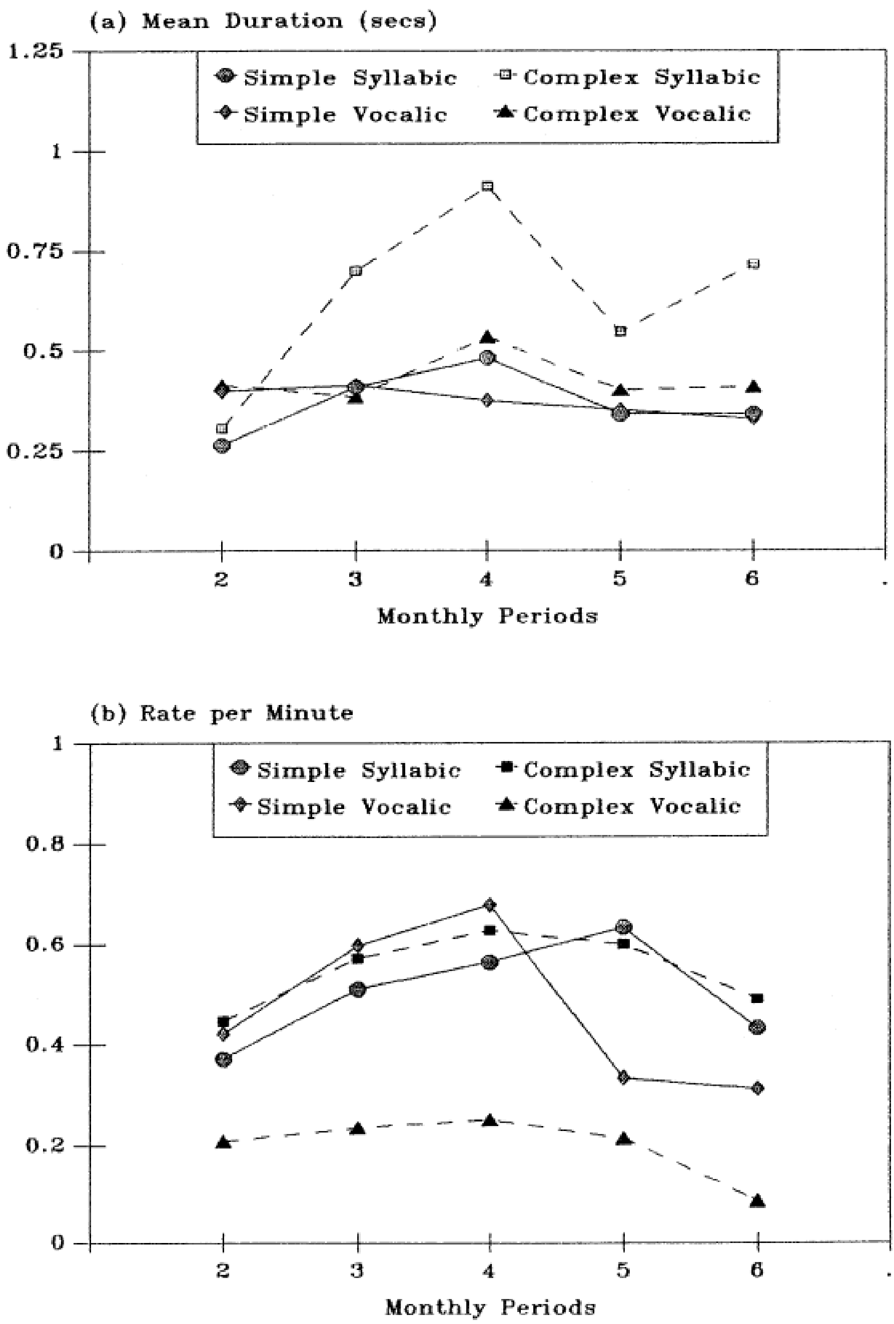

Figure 1. Developmental changes in infant non-distress vocalizations: (a) mean duration and (b) rate per minute.

complex melodic contours (see Table 3). The results suggested that there was a developmental change in the relation between speech quality and melodic complexity. A transitional period was found around the 4th month, before which vocalic sounds were more likely to be associated with simple melodic contours, and after which syllabic sounds were more likely to be associated with complex melodic contours.

DISCUSSION

This study examined developmental changes in infants' non-distress vocalizations during social interaction during the first 6 months of life. Results showed that there were significant differences in the mean duration and the rate per minute of infant non-distress vocalizations depending on their speech-like sound quality and complexity of the melodic contours. Syllabic (speech-like) sounds were found to be significantly longer and more frequent than vocalic (nonspeech-like) sounds. Infant vocalizations with complex melodic contours were longer but less frequent than those with simple melodic contours.

Specifically, complex syllabic sounds were the longest in duration, whereas complex vocalic sounds were the least frequent (rate per minute). These results replicate previous research findings which showed that, compared with vocalic sounds, syllabic sounds were longer and had more variable melodic intonations (Bloom, 1988, 1989, 1993; Masataka, 1995).

This study also found that the growth of infants' non-distress vocalizations during the first 6 months is characterized by non-linear developmental trajectories. While both cubic and quartic developmental trends (a dip after the peak and then a slight rebound) were found in the mean duration of complex syllabic sounds, a quadratic pattern (an inverted-U shape) was found in the rate of simple vocalic sounds. Regressive transitions in both types of vocalizations occurred around the 4th month. Previous findings have shown that the mean duration of infant vocalization increases over time (Delack, 1978; Laufer and Horii, 1976; Stark, 1989), which was not replicated in our study. The discrepancy may be because of the fact that different methods were used in previous studies including: (1) no differentiations were made for the types of non-distress vocalizations; (2) vocalizations analyzed in previous studies were collected from a variety of situational contexts (e.g. when alone, during social interaction, and during interaction with objects, etc.); (3) the duration of vocalization was derived spectrographically from the acoustic analysis during which vocalizations with short duration (<200 ms) were deleted; (4) only vocalizations with

Table 3. Frequency cross-tabulation of infant non-distress vocalization by speech quality by melodic complexity aggregated by monthly period

Melodic contour	Speech quality	Monthly periods					Total
		2	3	4	5	6	
Simple	Syllabic	64 (88.47)$^-$	106 (137.51)$^-$	109 (139.23)$^-$	124 (127.57)	86 (87.45)	489
	Vocalic	82 (57.15)$^+$	123 (88.82)$^+$	129 (89.93)$^+$	74 (82.40)	58 (56.48)	466
Complex	Syllabic	74 (68.28)	124 (106.12)	123 (107.44)	125 (98.45)$^+$	93 (67.48)$^+$	539
	Vocalic	38 (44.10)	48 (68.55)	45 (69.40)$^-$	49 (63 59)	18 (43.59)$^-$	198
Total		258	401	406	372	255	1692

Expected values are in parentheses.
Bonferroni adjustment was applied ($p \leq 0.0025$). '$-$': $z \leq 3.03$; '$+$': $z \geq 3.03$.

longer duration were analyzed; and (5) visual inspections as opposed to statistical tests were employed to demonstrate the pattern of developmental change. Nevertheless, despite methodological differences in data collection as well as classification and measurement of vocalization, the present study confirmed previous reports that the amount of infant vocalization produced in the context of social interaction accelerated first and dropped subsequently (Tipps *et al.*, 1981; Berger and Cunningham, 1983; Legerstee, 1991). The results from the current study suggest that the decline in the rate is primarily because of a reduction in the number of simple vocalic sounds.

The dissipation of simple vocalic sounds and the reduction in the duration of complex syllabic sounds after the 4th month demarcate a transitional period in infant early vocal development. Developmental reorganization in the relation between speech quality and melodic complexity found in this study further demonstrate the occurrence of a developmental shift in the infant vocal system. Before 4 months of age, vocalic sounds were associated with simple melodic contours, whereas after 4 months of age, syllabic sounds were associated with complex melodic contours. The onset of the transition around the 4th month coincides with the milestones depicted in stage models, e.g. the expansion stage in Oller's (1980) model, the vocal play stage in Stark's (1980, 1986) model, or the velar stage in Roug *et al.*'s (1989) model. Implied by structural stage theories, vocal development is a progressive invariant sequence characterized by abrupt qualitative augmentation (cf. Flavell, 1971). Nevertheless, results from the present study suggest that regressive growth may also characterize early vocal development. The non-linear developmental changes in the quantitative parameters of infant nondistress vocalization highlight that early vocal development is not merely progressive additions or accumulations. Rather, it also involves regressions (Vihman, 1996).

A question that needs further explanation is why the developmental transition occurs around the 4th month of life. A biological or neuromotor maturational theory is traditionally utilized to explain early vocal development. Early vocal production is conceived as an interaction of anatomical structures, and motor control of speech articulatory movement, which determine the input–output relation between articulatory movement and acoustic properties of vocalizations (Kent, 1981). Specific acoustic properties in vocal output are generated by certain speech articulatory movements (coordination between tongue, lips, and soft palate, etc.; Kent, 1981), which set boundaries for acoustic features in the vocal signals. Significant changes in the anatomical–physiological structure of the infant's vocal tract occur at around 4–6 months. There is a noticeable separation between the oral and the nasal cavities (Sasaki *et al.*, 1977, cited in Kent and Murray, 1982). As the anatomical systems mature, an increment in the control of speech articulatory movements allows the infant to produce non-nasal sounds (Kent, 1981). When speech motor movements by infants are well coordinated in their speed and timing, the acoustic features of the vocal output emulate more mature speech (such as syllabic sounds), which is characterized by longer duration with more prosodic fluctuations (Lieberman, 1986). The maturing anatomical structure and increasing neural-motor coordination ability are expected to enhance the production of speech-like (syllabic) sounds—becoming longer in duration and more complicated in melodic patterns over time.

Furthermore, according to the maturational model, while the exhibition of vocalic (non-speech-like) sounds is expected to decline over time, the production of syllabic (speech-like) sounds should increase. As implicated by the maturational model, the rate of simple vocalic sounds reaches its peak at around 4 months, and then decreases. However, the rate of syllabic sounds did not increase as this model suggested. In addition, the hypothesized elongation in the mean duration of syllabic sounds as the vocal production system matures was not observed in the present study. In fact, the mean duration of complex syllabic sounds was found to decrease around the 4th month. An alternative model is needed to explain the observed phenomena.

A dynamic systems perspective has been employed to explain various domains of early development such as motor skills (e.g. Thelen and Ulrich, 1991), facial expressions (e.g. Cameras, 1992), and communicative actions (e.g. Fogel and Thelen, 1987). This theoretical framework may provide useful insights into the regressive patterns found in infant early vocal development. According to a dynamic systems perspective, a vocal system consists of multiple components with high dimensionality. The coordination of over 70 muscles and many different body parts, ranging from the diaphragm to the lips, is required to produce an utterance. Infant vocalizations are the outcome of the emergent coordination of these different components (cf. Kent, 1984; Thelen, 1991), which not only drives but also constrains the development of the vocal system. During the process of self-organization in a complex system like vocal production, we might expect that only a small number of 'preferential linkages' (Fogel and Thelen, 1987) or 'attractor states' (Thelen and Ulrich, 1991) would be formed. As a result, only certain combinations of speech quality and melodic contour would occur. The changing patterns of association and dissociation between speech quality and melodic complexity found in this study may be the attractor states generated by the constraints of immature anatomical and neurophysiological structures. Because attractors are not rigidly structured, but are dynamically stable with flexibility, the preferential pattern of infant vocalization dissolves as the constraining conditions change over time. Therefore, a plausible interpretation for the regression may be that the progression is inhibited by the changes in the articulatory–acoustic relation. For certain articulatory configurations, a small change in articulatory characteristics may result in a large change in the acoustic properties in the vocal output (Kent, 1981). It can be speculated that before the acquisition of appropriate speed and precise sequencing of speech motor control, when infants produce complex syllabic sounds, the duration of the vocalization may become shorter before it is elongated. The driving force (or control parameters) of the 4-month shift, of course, is not limited to the anatomic and neurophysiological changes. The plausible contributing factors also include the developmental changes in infants' communicative, expressive, social, cognitive, perceptual, and motor repertories.[2]

According to the results of the present study, infants become more efficient in producing sounds with speech-quality by reducing non-speech-like sounds, which is demonstrated by the decrement in the rate of vocalic sounds with simple melodic contour after 4 months of age. From a dynamic systems perspective, regressive behaviors may be a necessary precursor for a system to move to the next level of development. In other words, regressions may be the surface effect of some underlining restructuring processes (cf. van Geert, 1994).

The reliability and validity of stage models of infant vocal development have been questioned by a number of researchers (e.g. Kent, 1984; Kent and Murray, 1982). The findings of our study also challenge the stage model account of development and further suggest that: (1) early vocal development may be a continuous quantitative changing process as opposed to a discontinuous qualitative stepwise shift; (2) the process may not be characterized by a linear increment but by a non-linear pattern involving both progressive and regressive changes; and (3) the change may be typified by the reorganization of dynamically stable patterns as opposed to the unfolding of invariant successively ascending sequences (cf. Stark, 1989). The construct of stage may be better used as a metaphor to convey the ideas of developmental change but is not suitable for use as a strategy to investigate the processes of a developing system (cf. Brainerd, 1978).

Because of the relatively small sample size and the limitation of vocal sampling from only one context (i.e. mother–infant social interaction), we caution the readers about the generalizability of the results of this study. Furthermore, large individual differences are typically observed in early vocal development; the current study did not address this issue. Further empirical investigations of the individual differences in both the quantitative and the qualitative aspects of early vocal development are needed. Moreover, even though it has been speculated that the reorganization in infant vocal system is associated with the onset of early social interaction (Stark, 1989), the relation between speech quality and melodic complexity of infant non-distress vocalizations and the quality of face-to-face interaction were not examined in this study. The importance of context on early development is emphasized by a dynamic systems approach (cf. Thelen and Ulrich, 1991). The quality and the quantity of vocal development in the context of a changing mother–infant communication system warrant further examination.

Notes

1. See Stark (1975) for a comprehensive analysis of infant cry, discomfort, and vegetative vocalizations.
2. Thanks to an anonymous reviewer for the suggestions.

REFERENCES

Bateson MC. 1975. Mother–infant exchanges: the epigenesis of conversational interaction. *Annals of the New York Academy of Sciences* **263**: 101–113.
Beaumont SL, Bloom K. 1993. Adults' attributions of intentionality to vocalizing infants. *First Language* **13**: 235–247.
Berger J, Cunningham CC. 1983. Development of early vocal behaviors and interactions in Down's Syndrome and nonhandicapped infant–mother pairs. *Developmental Psychology* **19**: 322–331.
Bloom K. 1988. Quality of adult vocalizations affects the quality of infant vocalizations. *Journal of Child Language* **15**: 469–480.
Bloom K. 1989. Duration of early sounds. *Infant Behavior and Development* **12**: 245–250.
Bloom K. 1993. Infant voice and mouth movement affect adult attributions. *Journal of Language and Social Psychology* **12**: 269–287.
Bloom K, D'Odorico L, Beaumont S. 1993. Adult preferences for syllabic vocalizations: generalizations to parity and native language. *Infant Behavior and Development* **16**: 109–120.
Bloom K, Russell A, Wassenberg K. 1987. Turn taking affects the quality of infant vocalizations. *Journal of Child Language* **14**: 211–227.
Brainerd CJ. 1978. The stage question in cognitive-developmental theory. *Behavior and Brain Sciences* **1**: 182–183.
Cameras LA. 1992. Expressive development and basic emotions. *Cognition and Emotion* **6**: 269–283.
Delack JB. 1978. Aspects of infant speech development in the first year of life. *Canadian Journal of Linguistics* **21**: 17–37.
Delack JB, Fowlow PJ. 1978. The ontogenesis of differential vocalization: development of prosodic contrastivity during the first year of life. In *The Development of Communication*, Waterson N, Catherine S (eds). Wiley: New York; 93–110.

D'Odorico L. 1984. Non-segmental features in prelinguistic communications: an analysis of some types of infant cry and non-cry vocalizations. *Journal of Child Language* **11**: 17–27.
Flavell JH. 1971. Stage-related properties of cognitive development. *Cognitive Psychology* **2**: 421–453.
Fogel A, Thelen E. 1987. The development of expressive and communicative action in the first year: reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology* **23**: 747–761.
Kent RD. 1981. Articulatory–acoustic perspectives on speech development. In *Language Behavior in Infancy and Early Childhood*, Stark RE (ed.). Elsevier North Holland: New York; 105–126.
Kent RD. 1984. Psychobiology of speech movement: coemergence of language and a movement system. *American Journal of Physiology* **246**: 888–894.
Kent RD, Murray AD. 1982. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of Acoustic Society of America* **72**: 353–365.
Kessen W, Levine J, Wendrich KA. 1979. The imitation of pitch in infants. *Infant Behavior and Development* **2**: 93–100.
Koopmans-van Beinum FJ, van der Stelt JM. 1986. Early stages in the development of speech movements. In *Precursors of Early Speech*, Lindblom B, Zetterstrom R (eds). Stockton: New York; 37–50.
Laufer MZ. 1980. Temporal regularity in prespeech. In *Infant Communication: Cry and Early Speech*, Murry T, Murray J (eds). College Hill: Houston, TX; 284–309.
Laufer MZ, Horii Y. 1976. Fundamental frequency characteristics of infant non-distress vocalization during the first twenty-four weeks. *Journal of Child Language* **4**: 171–184.
Legerstee M. 1991. Changes in the quality of infant sounds as a function of social and nonsocial stimulation. *First Language* **11**: 327–343.
Lieberman P. 1986. The acquisition of intonation by infants: physiology and neural control. In *Intonation in Discourse*, Johns-Lewis C (ed.). College-Hill: San Diego, CA; 239–257.
Lienert GA, Krauth J. 1975. Configural frequency analysis as a statistical tool for defining types. *Educational Psychology Measurements* **35**: 231–238.
MacNeilage PF, Davis BL. 1992. Motor explanations of babbling and early speech patterns. In *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, de Boysson-Bardies B, de Schonen S, Jusczyk PW, MacNeilage P, Norton J (eds). Kluwer: Netherlands; 341–352.
Marcos H. 1987. Communicative functions of pitch range and pitch direction in infants. *Journal of Child Language* **14**: 255–268.
Masataka N. 1995. The relation between index-finger extension and the acoustic quality of cooing in three-month-old infants. *Journal of Child Language* **22**: 247–257.
Masataka N, Bloom K. 1994. Acoustic properties that determine adults' preferences for 3-month-old infant vocalizations. *Infant Behavior and Development* **17**: 461–464.
Maskarinec AS, Cairns GF, Butterfield EC, Weamer DK. 1981. Longitudinal observations of individual infant's vocalizations. *Journal of Speech and Hearing Disorders* **46**: 267–273.
McCune L, Vihman MM, Rough-Hellichius L, Delery DB, Gogate L. 1996. Grunt communication in human infants (Homo sapiens). *Journal of Comparative Psychology* **110**: 27–37.
Nakazima S. 1975. Phonemicization and symbolization in language development. In *Foundations of Language Development*, vol. 1, Lenneberg EH, Lenneberg E (eds). Academic Press: New York; 181–187.
Nwokah EE, Hsu H, Dobrowolska O, Fogel A. 1994. The development of laughter in mother–infant communication: timing parameters and temporal sequences. *Infant Behavior and Development* **17**: 23–35.
Oller DK. 1978. Infant vocalization and the development of speech. *Allied Health and Behavioral Science* **1**: 523–549.
Oller DK. 1980. The emergence of the sounds of speech in infancy. In *Child Phonology: Production*, vol. 1, Yeni-Komshian G, Kavangh JF, Ferguson CA (eds). Academic Press: New York; 93–112.
Oller DK. 1986. Metaphonology and infant vocalizations. In *Precursors of Early Speech*, Lindblom B, Zetterstrom R (eds). Stockton: New York; 21–35.
Oller DK, Eilers RE, Steffens ML, Lynch MP, Urbano R. 1994. Speech-like vocalizations in infancy: an evaluation of potential risk factors. *Journal of Child Language* **21**: 33–58.

Oller DK, Lynch MP. 1992. Infant vocalizations and innovations in infraphonology: toward a broader theory of development and disorders. In *Phonological Development: Models*, *Research*, *Implications*, Ferguson CF, Menn L, Gammon CS (eds). York: Timonium, MD; 509–536.
Papousek M. 1992. Early ontogeny of vocal communication in parent–infant interactions. In *Nonverbal Vocal Communication: Comparative and Developmental Approaches*, Papousek H, Jurgens U, Papousek M (eds). Cambridge University Press: New York; 230–261.
Papousek M, Papousek H. 1981. Musical elements in the infant's vocalization: their significance for communication, cognition, and creativity. In *Advances in Infancy Research*, vol. 1, Lipsitt L (ed.). Ablex: Norwood, NJ; 163–224.
Papousek M, Papousek H. 1989. Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Language* **9**: 137–158.
Prescott R. 1975. Infant cry sound: developmental features. *Journal of Acoustic Society of America* **57**: 1186–1191.
Roug L, Landberg I, Lundberg LJ. 1989. Phonetic development in early infancy: a study of four Swedish children during the first 18 months of life. *Journal of Child Language* **16**: 19–40.
Sasaki CT, Levine PA, Laitman JT, Crelin ES. 1977. Postnatal descent of the epiglottis in man. *Archives of Otolaryngology* **103**: 169–171.
Scherer KR. 1982. The assessment of vocal expression in infants and children. In *Measuring Emotions in Infants and Children*, Izard CE (ed.). Cambridge University Press: Cambridge; 127–163.
Stark RE, Rose SN, McLiagan M. 1975. Features of infant sounds: the first eight weeks of life. *Journal of Child Language* **2**: 205–221.
Stark RE. 1978. Features of infant sounds: the emergence of cooing. *Journal of Child Language* **5**: 379–390.
Stark RE. 1980. Stages of speech development in the first year of life. In *Child Phonology: Production*, vol. 1, Yeni-Komshian G, Kavanagh JF, Ferguson CA (eds). Academic Press: New York; 73–95.
Stark RE. 1981. Infant vocalization: a comprehensive view. *Infant Mental Health Journal* **2**: 118–128.
Stark RE. 1986. Prespeech segmental feature development. In *Studies in Language Acquisition*, Fletcher P, Garman M (eds). Cambridge University Press: Cambridge, MA; 149–173.
Stark RE. 1989. Temporal patterning of cry and non-cry sounds in the first eight months of life. *First Language* **9**: 107–136.
Stark RE, Bernstein LE, Demorest ME. 1993. Vocal communication in the first 18 months of life. *Journal of Speech and Hearing Research* **36**: 548–558.
Thelen E. 1991. Motor aspects of emergent speech: a dynamic approach. In *Biological and Behavioral Determinants of Language Development*, Krasnegor NA, Rumbaugh DM, Schiefelbusch RL, Studdert-Kenndy M (eds). Lawrence Erlbaum: Hillsdale, NJ; 339–362.
Thelen E, Ulrich BD. 1991. Hidden skills. Monographs of the Society for Research in Child Development, 56 (Serial No. 223).
Tipps ST, Mira MP, Cairns GF. 1981. Concurrent tracking of infant motor and speech development. *Genetic Psychology Monographs* **104**: 303–324.
Tonkova-Yampol'skaya RV. 1973. Development of speech intonation in infants during the first two years of life. In *Studies of Child Language Development*, Ferguson CA, Slobin DI (eds). Holt, Rinehart, and Winston: New York; 128–138.
Trehub SE, Bull D, Thorpe LA. 1984. Infants' perception of melodies: the role of melodic contour. *Child Development* **55**: 821–830.
van Geert P. 1994. *Dynamic Systems of Development: Change between Complexity and Chaos*. Harvester Wheatsheaf: Hertfordshire, UK.
Vihman MM. 1996. *Phonological Development: The Origins of Language in the Child*. Blackwell: Oxford, UK.
von Eye A. 1990. *Introduction to Configural Frequency Analysis: The Search for Types and Antitypes in Cross-classifications*. Cambridge University Press: New York.

