Proceedings of DiSS-LPSS Joint Workshop 2010,
25-26 September 2010, Tokyo, Japan, pp. 93-96.

Pitch Patterns in the Vocalization of a 3-month-old Taiwanese Infant

Pei-Yu Hsieh

Graduate Institute of Linguistics, National Chung Cheng University, Taiwan
d96130004@ccu.edu.tw

Abstract

This paper studied pitch contours of a Taiwanese-acquiring infant at gooing stage. Breath group theory has shown that pitch patterns of this stage were physiologically-based [6]. Fall was expected to occur at the boundary of a breath group. It predicted that Fall to be the most common pitch contour, and the second high was Rise-Fall. But previous studies [8], [9] showed that Rise-Fall occurred more. We investigated patterns of an infant from six weeks old to twelve weeks old. Mean *f0* of basic contours of this stage were also shown. The *f0* range of Level, Fall, and Rise were reported. Our results showed four types of contours (Level, Fall, Rise, Rise-Fall) appearing at this stage. Consistent with the hypothesis, Fall was found to be most common. Rise-Fall was found to be the second high. Fall and Rise-Fall made up to almost seventy percent. Level contour was found to be rare. The mean *f0* of the infant at 3-month old was 400 Hz, higher than that of a toddler at 1;3 (370 Hz) and that of an adult (220 Hz). . The *f0* range was 700 Hz, greater than that of a toddler at 1;3 (450 Hz), and an adult (300 Hz).

Index Terms: vocalization**,** pitch, acquisition

1. Introduction

This paper investigated the pitch contour pattern in the production of a 3-month-old infant, whose ambient language is Taiwanese, one of Chinese dialects. Vocalization refers to voluntary sounds made by infants at the pre-linguistic stage [1]. Infants at prelinguistic stage do the vocal play to produce consonantal features, sound levels (murmurs or howls) and pitch contours [1]. Pitch is the first aspect of language children acquire, earlier than segments [2]. Pitch contours in early vocalization will develop into intonation in verbal stages. The linguistic usage of pitch develops into the complex prosodic systems [3]. In stressed languages, pitch is used to differentiate functional sentence-types, e.g. statement, yes/no question [4]. In tone languages, it represents tonal contrast in lexical meanings. Therefore, pitch development is crucial in understanding language acquisition. Basic contour were Level, Rise, and Fall. Complex contour were combinations of basic contours [5].

Breath-group theory proposed by [6] claimed Fall occurred more. It was physiologically-based. It was expected at the end of normal breath group due to the decrease of subglottal air pressure. Rise-Fall was also found to be common in infant speech, and it is also related to the subglottal air pressure change within a respiratory cycle [6].

Studies about pitch patterns of gooing stage were not many [7], [8], [9]. Fall was found to occur more frequently in infants at two-month old [7]. [8] studied 19 English-acquiring infants from one-month old to twelve-month old. Seven contours were found, including three basic contours and complex contours. Basic contours made up to twenty-five percent. It was found that Rise-Fall, Fall, and Rise were predominant. Level was rare (2%) during the first year. 21 infants at three-month old were investigated [9]. Basic contours and four other complex contours were found. In terms of the occurring frequency, Rise-Fall was the first high, Fall was the second high, and Level was the third [9]. Results of [8] and [9] were similar in 1) basic contours were found, 2) Rise-Fall was found to be the most frequent contour. Yet, The second point was not consistent with the hypothesis of breath-group theory, which is Fall was expected to be the most common. Besides, the methods used may be problematic. The complex contours defined by these two studies were not consistent. The traditional method used was based on perception and spectrograms. Contours defined may be different due to subjective judgment of individuals, and it influences the outcomes. Also, It is a longitudinal study that can show the developmental patterns each child follows over a period of time [10]. Cross-sectional studies [8], [9] assumed each individual pass through every developmental stage at the approximate time, but it was not true.

Therefore, the purpose of this study is to examine the pitch pattern of a 3-month-old Taiwanese-acquiring infant. The method used here was the adaptation of relational and independent analyses of intonational analysis [13], [14]. The data recorded were from as early as 6 weeks. As this longitudinal study continues, we may have a better understanding about how intonation develops into lexical tone, and whether there is a universal pattern due to the physiological constraints of infants..

2. Methodology

2.1. Child

One female infant, whose name was abbreviated HJ, growing up in a Taiwanese-speaking family, was recorded. The people taking care of the infant are Taiwanese speakers, including the parents and the nanny, and speak Taiwanese to the infant.

2.2. Recording procedure and data collection

The recording was made from the infant was 6-week old. The recording interval was about a week. The total recording time was 186 minutes in six recording sessions as shown in Table 1. Audio recordings were obtained when the infant was in a comfort state, talking to her dolls or caretakers. During the recordings, the caretakers try to elicit the infant's production by asking her questions, e.g. "What else do you want to say?" The recording time ranged from seven minutes to sixty-six minutes.

Table 1. *Data in this study*

	Age of the infant	Recording length (minute)
1	6 weeks	7 min.
2	7 weeks	7 min.
3	9 weeks	66 min.
4	10 weeks	51 min.
5	11 weeks	30 min.
6	12 weeks	25 min.
Total		186 mins

There are six recording sessions, totally 186 minutes. The recording of 8 week was missing, since the infant was sick, and had a sore throat at that week. She was not interested in the vocal play. The recording was not made during that week.

2.3. Recording equipment

The recording devices consisted of a Sharp MD-MS200 digital recorder, one high-quality microphone, and digital mini-discs. The mini-disc recordings were transferred into machine-readable wave files by using GoldWave version 5 at 16 bits with a sampling rate of 44100 Hz, in mono channel.

2.4. Acoustic analysis

The software for acoustic measuring was Praat [11]. The transferred long file was edited based on the following steps. Firstly, unnecessary data was excluded, for example, noise from the external environment, vocalizations overlapping with other speakers. Secondly, boundary of vocalizations was determined by 1 second or longer pause, and turn-taking [12]. Thirdly, Vegetative and reflexive sounds such as breathing noise, coughing, or hiccough were excluded [9]. Narrow-band (45 Hz) spectrograms were employed for the determination of *f0*. Following [9], two criteria were used to identify the *f0* contour of vocalizations. The first one was that they must have clear F1-F2 pattern; the second one was that they must be longer than 100 millisecond (ms), but shorter than 2 second. Under these criteria, 504 vocalizations were recognized as the data of this paper.

2.5. Pitch contour analysis

The method we used in analyzing pitch contour production was based on [13] and [14]. Their methods were the adaptation and application of relational analysis and independence analysis proposed by [10]. Relational analysis for segments compared children production with the adult correct production. For instance, if [ti] was produced for the target word /ki/ *key*, this would be analyzed as incorrect for /k/ [10]. Independence analysis did not compare children production with adult forms. Rather, if the child produced [ti] for /ki/, it showed that he/she was able to produce /t/, and /t/ was in his/her phonological inventory [10].

[13],[14] proposed relational analysis for pitch contour analysis referred to pitch change (the difference between the maximum and minimum *f0* of the pitch contour) in semitones [15], using the following formula: *[12/log(2)] *[log (max f0/min f0)]*. Independence analysis referred to the contour inventory. It consisted of sixteen different pitch contour (See Figure 1) adapted from [3] and [4]. The new contour needed to be produced at least twice, so it would be considered that the new contour was in child’s inventory. The classification of pitch contour analysis was based on the *f0* directionality, accent change (pitch change), and complexity. Directionality referred to the maximum *f0* and minimum *f0* of the prominent portion of contour. If the maximum *f0* preceded the minimum *f0*, it was considered Fall contour. If the minimum *f0* preceded the maximum *f0*, it was a Rise contour. Then, the second parameter of the pitch contour analysis was accent change (pitch change) by using the formula mentioned. Measures of 4 semitones for Fall and 3 semitones for Rises were used [13]. Complexity referred to the second directional change in contour, and the pitch change in semitones must be greater than 1 semitones. If it did not reach 1 semitone, the second directional change did not count, since it was not noticeable by perception [15]. The contours defined by [13] and [14] were shown in Figure 1.

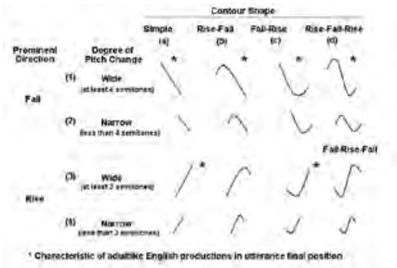

Figure 1: *Contour Inventory (Balog and Snow, 2007, p.125)*

Sixteen contours were shown in Figure 1, including simple Fall, simple Rise, and their combinations. They were used in this study in defining the pitch contour types produced by the infant.

3. Results & Discussion

This section reports the mean *f0* of the infant, the range of *f0*. We also compare the *f0* of the infant, a toddler at 1;3 [16], and an adult to see the variation of the *f0*. They are with the same ambient language, and Taiwanese-speakers.

The mean *f0* at three-month old was 400 Hz, ranging from 350 Hz to 450 Hz. [9] showed the average *f0* was 445 Hz, ranging from 350 to 500 Hz. The mean f0 of basic contours were displayed in Table 2. It was shown that the onset *f0* of Level and Fall was both around 450 Hz, and Fall decreased to 350 Hz. For Rise contour, the beginning *f0* was around 360 Hz, and increased to 420 Hz. Average *f0* contour of three basic contours was shown in Figure 2. Level was found to be slightly downward. Fall decreased from the onset to the offset. Rise increased along the *f0* contour.

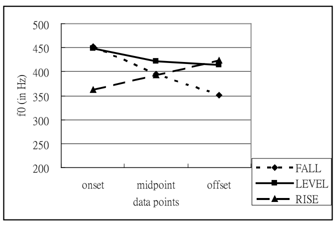

Figure 2: Mean f0 of basic contours

[16] studied a Taiwanese girl at 1;3 (MLU: 1.5 words). The average *f0* of the child was 370 Hz. The mean *f0* of level tone, falling tone, and rising tone was displayed in Table 3. The level tone and falling tone both begin around 410 Hz,

which was lower than that of the infant of this study. Falling tone decreased to 374 Hz. It was lower than the infant's production. The onset of rising tone was 365 Hz, midpoint 342 Hz, and offset 354 Hz. No explanation was given that rising tone did not increase along the *f0* contour.

The mean *f0* of a female Taiwanese was 220 Hz [17]. The adult's average *f0* was the lowest among these three, and the child at 0;3 had the highest mean *f0*. The vocal folds of the child were shorter and fewer mass, and it resulted in the high frequency [21]. The onset of level tone was around 250 Hz, and a little decreased along the *f0* contour. It was physiologically constrained. Falling tone began around 250 Hz, and decreased a lot. The midpoint value was 217 Hz, the offset was 193 Hz.

Table 2. *Mean f0 of basic contours (0;3)*

(in Hz)	onset	midpoint	offset
Level	448	421	413
Fall	453	395	351
Rise	362	392	424

Table 3. *Mean f0 of basic contours (1;3) (Lin, 2010)*

(in Hz)	onset	midpoint	offset
Level tone	410	417	425
Falling tone	411	404	374
Rising tone	365	342	354

Table 4. *Mean f0 of level, and falling tone (adult) (Hsieh, 2004)*

(in Hz)	onset	midpoint	offset
Level tone	249	242	235
Falling tone	250	217	193

The *f0* range of three people with the same mother tongue--the infant at 3-month old, a toddler at 1;3, and an adult was shown in Table 5. The *f0* range of the infant was 150 Hz to 900 Hz. The *f0* range of the toddler was from 150 Hz to 600 Hz [16]. The *f0* range of a female was from 150 Hz to 450 Hz [17]. The starting *f0* was all from 150 Hz, but the highest *f0* value was different. The infant was 900 Hz, the toddler was 450, and the adult was 300 Hz. The *f0* range of three people reduced with the growth of the age. The range of the infant was 750 Hz, the child 450, and the adult 300 Hz. It was due to the lengthening of vocal folds as the growth of age [21].

Table 5. *f0 range*

in Hz	Infant at 0;3	Child at 1;3	adult
Max *f0*	900	600	450
Min *f0*	150	150	150
f0 range	750	450	300

Table 6. *f0 range of Level (0;3)*

(in Hz)	onset	midpoint	offset
Max	891	890	901
Min	183	159	165
Range	707	731	737

Table 7. *f0 range of Fall (0;3)*

(in Hz)	onset	midpoint	offset
Max	909	700	512
Min	147	142	139
Range	761	558	373

Table 8. *f0 range of Rise(0;3)*

(in Hz)	onset	midpoint	offset
Max	727	741	800
Min	220	218	270
Range	507	523	530

The maximum and minimum *f0* of three basic contours were shown in Tables 6 to 8. The max *f0* of Level was around 900 Hz, and the min *f0* was around 200 Hz. The range was approximately 700 Hz. The mean *f0* of Fall was shown in Table 6. The maximum *f0* of each data point was 909 Hz, 700 Hz, and 512 Hz. The minimum *f0* was 147 Hz, 142 Hz, and 139 Hz. The *f0* range of each data point was 761 Hz, 558 Hz, and 373 Hz. The maximum and minimum *f0* of the average Rise was shown in Table 7. The max *f0* of each data point 727 Hz, 740 Hz, 800 Hz. The min *f0* was 220 Hz, 218 Hz, 270 Hz. The onset of Rise was lower than that of Fall and Level. Rise increased along the *f0* contour. The *f0* range of each data point was around 520 Hz.

The occurring frequency of contours was displayed in Table 9. Some observation can be made. Firstly, four contours appeared at this stage, including three basic contours and one complex contour. Secondly, basic contours took up to approximately seventy percent, as expected from vocal cord physiology [5]. Thirdly, complex contour was not common at this stage; there was only one type appearing- Rise-Fall. Fourth, regarding the occurring frequency, among totally 504 vocalizations, there were 205 tokens for Fall (41%). The appearance of Fall was physiologically-driven. The preference for Fall was found in the vocalizations of a three-month-old infant as claimed in [3], or even earlier- in the vocalizations of a six-week-old infant [18]. Falling contour was expected to be the most frequent based on Breath-group theory. It occurred at the end of breath-group boundary, due to the decrease of subglottal air pressure and the reduction of tension in vocal folds [6]. Fall was the easiest to produce and took the least effort to produce comparing with other contours, from physiology [19], [20].

Table 9. *Occurring Frequency of contours*

Contours	Fall	Rise-Fall	Rise	Level	Total
Tokens	205	137	108	54	504
%	41%	27%	21%	10%	

Rise-Fall was found to be the second high (27%) in terms of the occurring frequency. This was consistent with our second hypothesis—Rise-Fall contour was common in infancy, due to the subglottal air pressure change during a respiratory cycle [6]. Previous studies showed that Rise-Fall was also found to be common [8] [9], or even the most frequent. Subglottal air pressure change was raised to explain this phenomenon.

Rise was found to take up to twenty percent of the whole contours. The mechanism of producing a Rise was different from Fall [6]. Pitch-raising required more efforts [19]. Pitch-raising mechanism was complex. It was due to the increase tension in vocal folds, and the accompanying vocal fold length and thickness occur at the same time [5]. That explained why Rise was much less than Fall and Rise-Fall in the vocalizations of a three-month-old infant.

Among 540 tokens, Level (10%) was rare. Producing Level involved the sustaining of energy over times [5]. It may not be easy for infants due to physiologically constraints.

4. Conclusions

This study investigates the pitch pattern of a Taiwanese infant at 3-month old. Breath-group theory was used to analyze the patterns. The hypothesis of the theory was that 1) Fall was the most common; 2) Rise-Fall occurred a lot. The mean *f0* and the *f0* range of three basic contours was also reported.

The present finding shows the mean *f0* of a Taiwanese infant at 3-month old was 400. It was compared to that of a toddler at 1;3, and that of a female adult with the same ambient language. The average *f0* of the toddler was around 370 Hz, and that of an adult was 220 Hz. The mean *f0* decreased with age, due to the lengthening of the vocal folds [21].

The *f0* of three basic contours were also reported. The mean *f0* of three data points of Level was 448 Hz, 421 Hz, and 413 Hz. The mean *f0* of Fall was 453 Hz, 395 Hz, and 351 Hz. The mean *f0* of Rise was 362 Hz, 392 Hz, and 424 Hz. The *f0* range of the infant was greater than that of the toddler and the adult. The range of the infant was 750 Hz, the toddler 450 Hz, and the female 300 Hz. It was physiologically constrained.

Four types of pitch contour were found in this study, including Level, Fall, Rise, and Rise-Fall. Basic contour were found to be common and made up to seventy percent. Complex contour at gooing stage were few. As Breath-group theory predicted, Fall was found to be mostly common (41%) , and Rise-Fall was the second high (21%). Level was rare (10%). Fall was the easiest to produce and took the least effort [19]. The reason that Rise-Fall was found to occur more was due to the change of subglottal air pressure. Producing Level involves the energy-sustaining over a period of time [5], and it is not easy for an infant.

It would be very interesting as this longitudinal study continues to see whether the effect of ambient language or physiology takes the role affecting pitch patterns of babbling and early words.

5. Acknowledgements

Special thanks to Professor Jane Tsay for her insightful comments on this paper. Also, I am grateful to Professor David Snow for technical questions about pitch contour analysis.

6. References

[1] Boysson-Bardies, B. de, and DeBevoise, M.B. *How language comes to children: from Birth to two years.* The MIT Press. 1999.

[2] Lenneberg, E. *Biological foundations of language*. Malabar: R.E. Krieger. 1984.

[3] Crystal, D. "Prosodic development." In Paul Fletcher and Michael Garman (eds.) *Language acquisition*. Cambridge University Press., 174-197, 1986.

[4] Cruttenden, A. *Intonation*. Cambridge: Cambridge University Press. 1997.

[5] Zemlin, W. R. *Speech and Hearing Science: Anatomy & Physiology*. Prentice Hall. 1968.

[6] Liebrman, P. *Intonation, perception, and language*. Cabridge, MA: MIT Press. 1967.

[7] Stark, Rachel, E., Susan N. Rose and Margaret Mclagen. Features of infant sounds: the first eight weeks of life. *Journal of Child Language*, 2, 205-222, 1975.

[8] Delack, J.B. and Fowlow, P. J. "The ontogenesis of differential vocalizations: development of prosodic contrastivity during the first year of life." In N. Waterson and C. Snow (eds), *The development of communication*. New York: Wiley. 1978.

[9] Kent, R. D., & Murray, A. D. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*, 72, 353 – 365, 1982.

[10] Stoel-Gammon C. and C. Dunn. *Normal and Disordered Phonology in Children*. Pro-ed Publishing. 1985.

[11] Boersma, Paul & Weenink, D. *Praat: doing phonetics by computer* (Version 5.1.10) [Computer program]. Retrieved July 3, 2009, from http://www.praat.org/

[12] Branigan, G. Some reasons why successive single word utterances are not. *Journal of Child Language*, 6, 411 – 421. 1979.

[13] Balog, H. L. and Snow, D. "Relational and independent analyses of intonation in young children." Poster session presented at *the American Speech-Language-Hearing Association Convention*, Chicago, IL. 2003.

[14] Balog, H. L. and Snow, D. "The adaptation and application of relational and independent analyses for intonation production in young children. *Journal of Phonetics*, 35, 118-133. 2007.

[15] Burns. E. M. and Ward. W. D. "Intervals, scales, and tuning." In d. Deutsch (Ed.) *The psychology of music*, 241, 269, New York: Academic Press. 1982.

[16] Lin, Zhi-yuan. "*Acoustic analysis of tone production in a Taiwanese-speaking child from age of 1;2 to 1;4.*" (manuscript).

[17] Hsieh, Pei-Yu. *Variation in Taiwanese Tones.* Master Thesis. National Chung-Cheng University. 2004.

[18] Hsieh, Pei-Yu. "Pitch Patterns in the Vocalization of a 2-month-old Taiwanese Infant." *Paper presented at National Conference of Linguistics*, Taiwan. 2009.

[19] Ohala, J. and Ewan, Williams. 1973. Speed of Pitch change. [Abstract]. *Journal of Acoustic Society of America*, 53, 345.

[20] Ohala, J. 1978. Production of Tone. In V. A. Fromkin (ed) *Tone: A Linguistic Survey*. New York, Academic Press.

[21] Eckel HE, Koebke J, Sittel C, Sprinzl GM, Pototschnig C, Stennert E. 1999. Morphology of the human larynx during the first five years of life studied on whole organ serial sections. *Ann Otol Rhinol Laryngol*, 108:232–238