Intrinsic F0 of vowels in the babbling of 6-, 9-, and 12-month-old French- and English-learning infants

D. H. Whalen
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

Andrea G. Levitt
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511 and Department of French,
Wellesley College, Wellesley, Massachusetts 02181

Pai-Ling Hsiao
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511 and Department of Linguistics,
University of Connecticut, Storrs, Connecticut 06269

Iris Smorodinsky
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511 and Department of Linguistics,
Yale University, New Haven, Connecticut 06510

(Received 31 March 1994; accepted for publication 28 November 1994)

In every language so far examined, high vowels such as [i] and [u] tend to have higher fundamental frequencies (F0s) than low vowels such as [a] and [æ]. This intrinsic F0 effect (IF0) has been found in the speech of children at various stages of development, except in the one previous study of babbling. The present study is based on a larger set of utterances from more subjects (six French- and six English-learning infants), at the ages 6, 9, and 12 months. It is found, instead, that IF0 appears even in babbling. There is no indication in these data of a developmental trend for the effect, and no indication of a difference due to the target language. These results support the claim that IF0 is an automatic consequence of producing vowels.

PACS numbers: 43.70.Ep, 43.70.Fq

INTRODUCTION

The relationship between vowel height and fundamental frequency (F0) has been noted for at least 60 years (Taylor, 1933). High vowels such as [i] and [u] tend to have higher F0's than low vowels such as [a] and [æ]. The mechanism for this "intrinsic F0" (IF0) or "intrinsic pitch" has been the subject of great dispute (see the reviews in Ohala and Eukel, 1987; Sapir, 1989; Fischer-Jörgensen, 1990), but the consistency of the effect is not in question (Whalen and Levitt, in press). Every language that has been examined for IF0 (31 are listed in that work) has been found to have it, and these languages represent 11 of the world's 29 major language families. IF0 has been found not only in languages such as English (Peterson and Barney, 1952) and French (D' Cristo, 1982) that use F0 primarily for stress and intonation, but also in tone languages such as Mandarin (Shi and Zhang, 1987) that use F0 changes to distinguish words. IF0 seems to be insensitive to the size of the vowel inventory as well, since both small (e.g., Japanese with 5 vowels) and large (e.g., German with 14 systems show similar effects (Whalen and Levitt, in press).

With such universality, IF0 has typically been assumed to be an automatic consequence of vowel articulation. Indeed, the theories reviewed in Sapir (1989) take this as a given. Under that assumption, it is of great interest whether the vowels of babbling will show this effect, since the babbling child presumably has no vowel categories per se, but simply vocalic articulations. If the child's vocal apparatus already has the interconnections that produce the IF0 effect in adults, then we should see IF0 in babbling. If there are significant anatomical or coordinative differences between infants and adults, perhaps IF0 will not appear in babbling.

Only one study that we have found that examined this question (Bauer, 1988) examined two infants at 9 and 13 months. There were 201 vowels measured at 13 months and an unreported number for the earlier age. Vowels were put into one of four broad classifications: high front, high back, low front, or low back. Bauer found no effect of height, but did find an effect of front/back. He attributed this to the high position of the larynx in the infant (Crelin, 1987). This high position also leads to a more vertical orientation, which might lead to more influence of the tongue pulling in the front/back dimension.

As a note of caution, though, the number of subjects and the number of tokens in Bauer's study were both rather small. The size of the study can greatly affect the outcome, as can be seen in a similar failure to find a vowel height effect, this time in running speech. Umeda (1981) measured approximately 200 vowels from two speakers in spontaneous conversation. She found no evidence of IF0 and concluded that it was not present in running speech. However, there are a great many factors that influence F0 in speech, and these were not controlled for in her study. To counteract this variability, it is necessary either to increase the number of observations, or to control the context. When factors such as sentence focus and segmental environment are properly controlled, even running speech shows the effect (Ladd and Silverman, 1984; Shadle, 1985). We can presume, then, that a larger sample of unrestricted text would show the effect.
<table>
<thead>
<tr>
<th>Study</th>
<th>Age (yrs)</th>
<th>N</th>
<th>Sex</th>
<th>F0 for [i]/u</th>
<th>F0 for [a]/e</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiCristo, 1982</td>
<td>adult</td>
<td>1</td>
<td>fem</td>
<td>239</td>
<td>226</td>
<td>13(5.8)</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>3</td>
<td>male</td>
<td>133</td>
<td>124</td>
<td>9(7.3)</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>33</td>
<td>male</td>
<td>139</td>
<td>126</td>
<td>13(10.3)</td>
</tr>
<tr>
<td></td>
<td>adult</td>
<td>28</td>
<td>fem</td>
<td>233</td>
<td>211</td>
<td>22(10.4)</td>
</tr>
<tr>
<td>Petronio and Barney, 1952</td>
<td>adult</td>
<td>15</td>
<td>both</td>
<td>274</td>
<td>253</td>
<td>21(8.3)</td>
</tr>
<tr>
<td></td>
<td>&quot;child&quot;</td>
<td>3</td>
<td>both</td>
<td>294</td>
<td>262</td>
<td>32(12.2)</td>
</tr>
<tr>
<td>Petronio, 1961</td>
<td>5–11</td>
<td>97</td>
<td>both</td>
<td>250</td>
<td>229</td>
<td>21(9.2)</td>
</tr>
<tr>
<td>Glaze et al., 1990</td>
<td>10–12</td>
<td>46</td>
<td>both</td>
<td>248</td>
<td>229</td>
<td>19(8.3)</td>
</tr>
<tr>
<td>Hillenbrand et al., in press</td>
<td>6–10</td>
<td>30</td>
<td>both</td>
<td>287</td>
<td>267</td>
<td>20(7.5)</td>
</tr>
</tbody>
</table>

And, of course, it is not possible with babbling to restrict the context, so an increase in sample size is our only alternative. Thus the issue of $IF0$ in babbling cannot be considered to be settled, and the present study attempts to increase our understanding of this issue.

Although most researchers assume that $IF0$ is an automatic consequence of vowel production, others hold that $IF0$ is a deliberate enhancement of the speech signal by the speaker (Diehl and Kluender, 1989; Diehl, 1991). This account assumes that the perception of vowel height is a function not only of $F1$ frequency but of the difference between $F0$ and $F1$ (Traumhüller, 1981) and that speakers intentionally increase their $F0$ with high vowels to make this difference larger that it would otherwise have been. The universality of $IF0$, on this account, only argues for the usefulness of this particular enhancement. In babbling, however, there is no communicative intent and thus no distinctions to enhance. So the enhancement account should predict that $IF0$ will not appear in babbling. Even Bauer’s (1988) finding, if it is correct, would be inconsistent with the enhancement account, since it implies an automatic (though different) mechanism for $IF0$. If $IF0$ were not found for babbling, the enhancement account would seem to be supported, with the assumption that $IF0$ would be an enhancement acquired later in development.

If $IF0$ is found for babbling, the most likely explanation is that it is not only universal but automatic. For the enhancement account to accommodate such a result, it would seem that an inductive explanation would be necessary. That is, since children hear this vowel height/$F0$ correlation in whatever adult language they hear, they include it in their babbling. Enhancement per se should not be an issue, since there are (presumably) no categories to enhance, but the imitation might be complex enough to include small $F0$ changes. This issue will be addressed further in Sec. III.

If enhancement is operative, we might expect there to be a developmental trend toward increased usage of the enhancement. Previous studies of $IF0$ in older children, with ages ranging from 5 to 11 years, show no indication of a developmental trend in $IF0$. Table I presents results from five published studies (Peterson and Barney, 1952; Peterson, 1961; Sorenson, 1989; Glaze et al., 1990; Hillenbrand et al., in press). We have averaged the two high vowels [i] and [u] and the two low vowels [a] and [æ]. As can be seen in the difference column of Table I, there is variability, especially in the Sorenson values where the $N$ was small (only 3 per cell). But there is no indication of an overall trend toward larger (or smaller) effects.

If $IF0$ is universal, then we would expect to find similar patterns in the babbling of infants from any language environment. If $IF0$ is deliberate enhancement, we might expect that different languages would use the enhancement to different degrees. This difference might then appear as a difference in the babbling behavior of children in different language communities. The present study takes a first step in assessing the universality of $IF0$ in babbling by examining infants in two language environments, English and French. We have already found intonational differences between these two language groups in an earlier study (Whalen et al., 1991). That study included 10 of the 12 subjects analyzed here. Since these children are using $F0$ in different ways in their babbling, it is certainly possible that they would treat $IF0$ differently if they were producing $IF0$ deliberately.

$IF0$ in babbling, then, needs further examination. The present study examines the babbling of 12 infants, 6 each in English and French environments. The infants were recorded in the home at 6, 9, and 12 months of age. We measured all the vowels except for the central (e.g., [ə]) and lower-mid (e.g., [ɛ] and [ɔ]). This resulted in 7325 tokens to analyze. With a larger set of results, we can more confidently address
the issue of whether IF0 is automatic or under the speaker’s control.

I. METHODS

A. Subjects

The speakers were 12 infants, 6 learning French as their native language and 6 learning American English. The French infants were all living in Paris or its environs. The American infants lived in various cities on the northeast coast of the United States.

B. Stimuli

The utterances for the present study were selected from recordings made at weekly intervals by the parents of the children. Each infant was recorded in the home on a cassette tape recorder (Panasonic RQ 3145 or Marantz PMD 430) using a high quality microphone (Realistic supercardioid 33992A). Individual recording sessions lasted approximately 10–20 min. The parents were asked to choose a time when the child was likely to be alert and unlikely to cry. As far as possible, the microphone was held 20 cm from the baby. If necessary, the parent could attempt to induce babbling by speaking to the child (stopping, of course, when the infant began vocalizing). Additional comments about the session were recorded by the parent on a form provided with each tape.

All utterances from the 6-, 9-, and 12-month tapes were digitized onto the Haskins Laboratories VAX computer system. They were low-pass filtered at 9.6 kHz and sampled at 20 kHz, with preemphasis (Whalen et al., 1990). We excluded cries, whispers, and various vegetative sounds. If an utterance contained a combination of speech and nonspeech, we would try to transcribe the speech.

All the utterances were transcribed by the third author, a native speaker of Mandarin Chinese. He is phonetically trained, and has experience with a wide variety of languages. Transcriptions were made from the digitized waveform, with the help of either of the Haskins Laboratories program HADES (Rubin, 1995) or SIGNALIZE® (Keller, 1990). With these programs, the whole utterance could be heard repeatedly, as could any selected portion of the utterance. The overall character of the waveform also gave indications of possible syllables. The symbols of the International Phonetic Alphabet were used, with the understanding that some of the utterances would be very difficult to transcribe. We felt that obtaining more detailed transcription was worth the effort involved, since this allows us to make more comparisons than Bauer’s (1988) four-way classification.

Once the transcriptions were made, we selected the following vowels for analysis: high front ([i y i ɪ]), midfront ([e ø]), low front ([æ a ə]), high back ([u u ʊ]), midback ([o ø]), and low back ([a ø ə]). (Strictly speaking, [æ] is low central, but there were few enough members of this group anyway, so we included it.) While we did want to have the most accurate transcription possible, it was not possible to analyze the results any more finely than this, primarily because of the small number of instances of many of the vowels. There were a handful of tokens that were nasalized; these were simply included without any indication of the nasalization. We also treated all vowels without regard to their consonantal environment.

C. Analysis

All fundamental frequencies were measured from the speech waveform, by hand, using either HADES or SIGNALIZE®. The following procedure was used: For each syllable containing one of the vowels of interest, the main period of vowel activity was delimited. Then a location 40% of the way into this segment was found. In the best case, we would then measure five pitch periods to the left of the point and five to the right. The duration of this ten pitch period segment was then translated into an average F0 for that measurement point. In some cases, the pitch periods immediately around the 40% point were not measurable, either because the waveform was noisy or low in amplitude, or otherwise unclear. In those cases, the nearest ten measurable pitch periods within that syllable were chosen. Some tokens that had been transcribed proved to be too noisy or too faint to measure. Table II presents the number of measured tokens for the 12 subjects at the three ages. Two of the subjects lacked recordings at some of the months: JZ was missing the 9- and 12-month recordings, and MB was missing the 12-month recording. Both were French subjects. Another French subject, YC, lacked 12-month recordings but had 11-month ones. The 11-month recordings were used for the 12-month data for her. Two other subjects had sparse data at 1 or 2 months, so these were supplemented with recordings from an adjacent month. For English subject MA, 25.8% of the 6-month data was from the 6-month recordings, while the remaining 74.2% came from the 7-month recordings. Also for this subject, 37.7% of the 12-month data was from the 12-month recording, while the remaining 62.3% was from the 11-month recordings. Finally, for French subject MB, 35.0% of the 6-month data came from the 6-month recordings, while the remaining 65.0% came from the 7-month recordings.
TABLE III. F0 values for the six vowel categories for the 12 subjects.

<table>
<thead>
<tr>
<th></th>
<th>Front Mean F0</th>
<th>N</th>
<th>% of total</th>
<th>Back Mean F0</th>
<th>N</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>405.7</td>
<td>519</td>
<td>7.1</td>
<td>381.8</td>
<td>302</td>
<td>4.1</td>
</tr>
<tr>
<td>Mid</td>
<td>364.6</td>
<td>5254</td>
<td>71.7</td>
<td>359.9</td>
<td>289</td>
<td>1.2</td>
</tr>
<tr>
<td>Low</td>
<td>332.2</td>
<td>508</td>
<td>11.0</td>
<td>330.6</td>
<td>353</td>
<td>4.8</td>
</tr>
</tbody>
</table>

All of the target vowels were measured, with an exception for one subject. American subject NG had a large corpus, but the vast majority of her vowels were [e]. At six months, [e] was approximately 45× as frequent as the next vowel. At 9 months, the ratio was around 17 to 1. By 12 months, [e] outnumbered its nearest rival by a mere factor of 10. In order both to keep the representations of the vowels relatively similar, and to cut down on the amount of work required for this subject, only selected [e]'s were analyzed for her. For each age, a number of [e]'s was counted out (45, 17, or 10, for the three ages). The utterance containing that [e] was analyzed for all its [e]'s. Thus if there was only one [e], then that would be the only one analyzed. If that utterance happened to have several [e]'s, all of them were analyzed. In this way, [e] was still the most frequent vowel, but only by an overall factor of 2.3.

The distribution of these vowels is similar to those found in previous studies. In one cross-language study (Boysson-Bardies et al., 1989), back vowels were found to be relatively rare (6.6% of the utterances in French and English), though the low back vowels were the most common of those. In the present study, by contrast, the high back vowels accounted for a higher proportion of the back vowels than was the case for the other study. (The proportion of front vowels overall would be higher if we had not excluded many of NG's [e] vowels.) Our proportions are more in agreement with Buh's (1980) one English-learning infant. The selection criteria used here were too different to allow a direct comparison with the de Boysson-Bardies et al. (1989) study, but the distribution of vowels analyzed here is at least qualitatively similar to that found in other studies.

F0's larger than 700 Hz were excluded from the analysis. These represented 4.3% of the 7651 tokens measured. Such extreme values, while common in babbling, distort the means for those cells with small N's. It is also possible that a different phonation type is involved in such high F0's, which would be a second reason to exclude these values. The selection process resulted in 7325 tokens being measured for the 12 subjects.

II. RESULTS

Means for the six vowel types for the 12 subjects are given in Table III. Also given are the number of tokens that went into each value. Table IV gives the size of the IF0 effect, both for height and front/back. The front/back difference is given as the front vowel mean minus the back vowel mean. In this way, any difference that matches the results of Bauer (1988) will be positive in value, while contrary results will be negative. As can be seen, there is a positive difference for height for 10 of the 12 subjects. For front/back, only 5 of the 12 subjects match Bauer's results.

For an analysis of variance, we operated on the means for each of the six cells for the 12 subjects. An analysis that used each observation was attempted, but the enormously large degree of freedom for the error term meant that almost any difference, however trivial, appeared significant. Using the means also gives the subjects with fewer productions a stronger say in the analysis. Since the differences among speakers, not tokens, are of primary importance, this result is to be desired.

The analysis, then, included the grouping variable language (English or French), and two within factors, height and front/back (with 3 and 2 levels, respectively). Three of the subjects (MM, NM, and MB) have missing cells, due to the lack of any instances of the mid back vowel category. Rather than reject these subjects from the analysis, these cells were replaced with the means of the five other cells for these subjects. This is a conservative approach to data replacement, since it will tend to minimize differences that actually exist. Language was not a significant main effect [F(1,10)<1, n.s.], indicating that the babblers had roughly equivalent overall F0's. Height was a significant factor [F(2,20)=16.62, p<0.001], while the interaction with language was not [F(2,20)=2.09, n.s.]. Front/back was also not a significant factor [F(1,10)<1, n.s.]; neither was the interaction with language [F(1,10)<1, n.s.]. The two-way interaction of height and front/back was significant [F(2,20)=4.48, p<0.05], but the three-way interaction with language was not [F(1,10)=1.09, n.s.].

For the analysis by age, it was necessary to restrict the number of cells. By the time we break the results down into the six categories and the three ages for the 12 subjects, 45 of the 216 cells are empty. Most of these are for the low back and mid back vowels. Therefore we analyzed the four other cells, as a single factor of vowel quality with four levels, so that only differences among the four cells can be tested, not the front/back and high/low dimensions. Eighteen of the 45 empty cells come from the two subjects who lacked certain

TABLE IV. Size of the high/low difference and the front/back difference in F0 for the 12 subjects. The first and third rows are in hertz and the middle row is the high/low difference expressed as a percentage of the low vowel F0.

<table>
<thead>
<tr>
<th></th>
<th>French</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initials</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>NM</td>
</tr>
<tr>
<td>High-low</td>
<td>56.0</td>
<td>94.3</td>
</tr>
<tr>
<td>(as %)</td>
<td>16.8</td>
<td>23.9</td>
</tr>
<tr>
<td>Front-back</td>
<td>-11.1</td>
<td>-19.5</td>
</tr>
</tbody>
</table>

Whalen et al.: Intrinsic F0 of vowels in babbling 2536
TABLE V. Mean F0’s for four of the six vowel categories for the ten subjects that had measurements at each of the months analyzed. The mid-back and low back categories were missing for many of the subjects for one month or another. The last row shows the difference between the mean of the two high vowel categories and the low vowels category.

<table>
<thead>
<tr>
<th>Age in months</th>
<th>6</th>
<th>9</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean F0</td>
<td>N</td>
<td>Mean F0</td>
</tr>
<tr>
<td>High back</td>
<td>375.5</td>
<td>88</td>
<td>369.8</td>
</tr>
<tr>
<td>High front</td>
<td>402.3</td>
<td>115</td>
<td>412.5</td>
</tr>
<tr>
<td>Midfront</td>
<td>350.4</td>
<td>1757</td>
<td>381.8</td>
</tr>
<tr>
<td>Low front</td>
<td>313.6</td>
<td>227</td>
<td>336.7</td>
</tr>
<tr>
<td>Diff. between</td>
<td>743.3</td>
<td>54.5</td>
<td>61.9</td>
</tr>
</tbody>
</table>

months, as mentioned before: JZ had no 9- or 12-month data, and MB had no 12-month data. These two subjects were excluded from this analysis, so that the remaining subjects had no missing cells. The ANOVA factors were vowel, with four levels, language, with two levels (English and French), and age, with three levels (6, 9, and 12 months).

In this analysis, as before, language was not a significant factor \( F(1,8) < 1 \), n.s.). It did not enter into any significant interactions either. Age was not a significant main effect \( F(2,16) < 1 \), n.s.), which is to be expected: Even though F0 lowers throughout development (see Table I), the time elapsed here is too short to show this effect. Vowel is a significant main effect \( F(3,24) = 8.22, p < 0.001 \), again showing the height effect. The critical interaction, age by vowel, is not significant \( F(6,48) = 1.97, n.s.)

, giving no indication of a difference in the effect over the six months involved here (see Table V). Even if we analyze each month separately (despite the lack of an interaction), the IP0 effect is present at each age. The separate analyses are strong for the 6-month \( F(3,27) = 9.61, p < 0.001 \), and 12-month \( F(3,27) = 5.05, p < 0.01 \) measurements, and somewhat less robust for the 9-month \( F(3,27) = 2.94, p = 0.0512 \). There is no evidence of change in F0 over this time span.

Since we relied on our transcriptions to separate the vowels into categories, we need to be sure that we can do this independently of F0. In adult speech, it is certainly clear that different vowels can be produced with a wide range of F0’s without losing the vowel’s identity. With babbling, however, it is not possible to ask the speaker to reproduce a particular vowel. One way of avoiding the vowel identity problem would be to correlate F0 with F1. Since F1 is lower with the high vowels and higher with the low vowels, there should be a negative correlation between F0 and F1 when IF0 is present. In the Peterson and Barney (1952) data, in fact, there is such a correlation if we examine the three speaker groups (the 33 adult males, the 28 adult females, and the 15 children) separately. When we correlate each individual production (there were two per vowel) for each vowel for all the speakers, we obtain the following correlations: males, \( r = -0.16 (p < 0.001) \), females, \( r = -0.20 (p < 0.001) \), and children, \( r = -0.10 (p < 0.10) \). The correlation does not reach significance for the children either because of greater variability of their values or the smaller number of subjects.

When we examined our babbling data, however, there was a positive correlation between F0 and F1, but this was due to the fact that the formants were almost invariably excited by a single harmonic. With a mean F0 of 370 Hz, the formants in our set of babbles are poorly represented. If a harmonic happens to be at the center frequency of a formant, the two nearest harmonics would be approximately 15 db lower in amplitude even with a bandwidth of 100 Hz. [For adults, bandwidths typically remain in the range 50–60 Hz for formant values up to 2000 Hz (Dunn, 1961).] Harmonics with such low amplitudes are too close to the background level to contribute to the measurement of the formant. Occasionally in our measurements, we found two harmonics of equal amplitude, and it was possible to assume that the center frequency of the formant was between them. (With such limited measurements of the formant frequencies, it was, of course, impossible to measure the bandwidth with any confidence.) The appearance of two harmonics was uncommon, so the formant value was much more likely to be identical to one of the harmonics, resulting in the positive correlation between F0 and formant frequency.

As a final check on the possible misperception of F0 as vowel quality, we examined the distribution of the vowel categories by F0. If F0 were the only factor, then the distributions should be distinct. If any vowel can occur on any F0, then the distributions should be greatly overlapped. Figure 1 shows a highly overlapped pattern. For that figure, the number of tokens of a particular vowel in an F0 range or “bin” of approximately 16 Hz was counted. The top panel shows the proportion of all the vowels in the six categories. Because the midfront category is so disproportionately represented, it is hard to see the other distributions. Thus the lower panel shows the same data with a ceiling on the midfronts. As is clear, there are vowels of each category at every level of F0. Certainly the distributions are different, since that is what the IF0 effect consists of. But it is not the case that a high F0 was enough to cause a perception of a high vowel. The identifiability of the vowels was apparent throughout the F0 range. Thus there is no evidence of any large perceptual bias in the transcriptions.

It is impossible to rule out smaller perceptual biases which might have influenced the results. Indeed, small effects of F0 on the identification of ambiguous vowels have been found in one study by Reinhold Peterson (1986). Using synthetic vowels ranging from [u] to [o], he found that the most ambiguous vowel received more [u] responses with a high F0 compared with the low F0. The effects were quite small and never enough to change the majority decision. In addition, it was only possible to shift an ambiguous vowel from one category to a neighboring category. The results of Gottfried and Chew (1986), in which a wide range of F0's for sung vowels was used, also show extremely few instances in which the height of the perceived vowel differs by more than one level. Thus even if there were bias effects in the present transcriptions, such biases would not account for the F0 difference between the low vowels and the high vowels. Despite the impossibility of completely ruling out small bias effects, then, the pattern of results strongly suggests that the effects we have found are due to the vowel articulation and not to the transcription.
different qualities is random in this distribution. Thus it is very easy to have several utterances with high, even squeaky pitch with a low vowel. It takes a large sample for this to average out. As can be seen in Table IV, there were two subjects who showed higher F0's for low vowels, and they were two among those with the smallest number of tokens to analyze.

The differences between front and back vowels were not consistent from subject to subject in the present study. This is unlike Bauer's (1988) results but like the adult studies (Whalen and Levitt, in press). Given that Bauer's explanation of the front/back effect as due to the high larynx position is a plausible one, we need to explain why this high position does not change the F0 effect. In fact, the high position seen in Crelin's x-ray images is somewhat misleading, since most of those were taken at rest. As Crelin himself notes (1987: p. 96), the larynx is pulled down into a much more adultlike position during speech (and screaming). That is, infants must work to make their vocal tracts appear more adultlike, and doing so seems to bring their larynx into the same relationship with the tongue that adults have. Since they show the same F0 as adults, it seems likely that the same mechanism is involved as well. One might still suppose that different children might adopt different strategies, but even the two subjects who showed a contrary effect for height were inconsistent for front/back: Subject EC had the difference that Bauer found, while MB went in the other direction. So, as with the adults, there is no consistent effect of the front/back dimension on F0.

The present study also found no evidence of a developmental change over the 6-month span examined. This is consistent with the universality of IF0 (Whalen and Levitt, in press) and with the lack of any evidence of a developmental change later in life (Table I). The overall percentage of difference found for the babblers was 13.9% (as calculated from Table III). This is slightly higher than that found for other studies (Table I), but a statistical artifact is probably the cause. If we had transformed the measured F0's into a semitone scale before averaging, the effect of the very high F0's would have been reduced, and the difference between high and low vowels would probably have been much more similar, too. Certainly, if there is any developmental trend, it is for less F0 rather than more. This does not fit with the enhancement hypothesis.

The present results are at odds with the one previous study (Bauer, 1988) and call into question the explanation given there. The difference is most likely due to the difference in sample size (200+ tokens versus the present 7000+). In addition, his developmental change was based on a comparison of the size of two different F ratios, which is not a reliable method of comparing results across data sets. It is also risky to assume that the absence of a significant difference means that there is no effect. However, the measurements here are sufficiently strong to show the F0 effect at each of the three ages analyzed, so at least we know that the effect is not absent at any of the ages. It seems likeliest that there is no developmental change in F0.

These results also cast doubt on the description of F0 as a deliberate enhancement of the speech signal (Diehl and


ACKNOWLEDGMENTS

This research was supported by NIH Grant No. DC-00403 to Haskins Laboratories and Catherine Best. Portions of this research were presented at the 2nd Meeting of the Acoustical Society of America, Denver, Colorado, October 1993, and the 9th International Conference on Infant Studies, Paris, June 1994. Additional help with the stimuli was provided by Michele Sancier, Winifred McGowan, and Julia Irwin. We thank Arthur S. Abramson, Catherine T. Best, Hartmut Traunmüller, Keith Johnson, and an anonymous reviewer for helpful comments.