ON THE VOWEL AND ITS NATURE, BETWEEN EIGHTEEN MONTHS AND FIVE YEARS

John H. V. Gilbert, Phonetics Laboratory, Division of Audiology and Speech Sciences, University of British Columbia, Vancouver, Canada

Introduction

For a number of years, we have been interested in the development of vowels in children between eighteen months and approximately five years of age (chronological age, CA), and have conducted a number of studies which have been directed toward questions in both production and perception. In this paper I shall review them, and some other studies which relate to the title (Gilbert a-fl).

Our original curiosity about the development of vowels was motivated largely by two factors; the first being that in 1967, there was very little information relating to this particular aspect of phonological acquisition; the second being that the classic paper of Peterson and Barney (1952) tantalizingly showed marked differences in vowel formant measure between children and adults without at any place in the paper stating how old the subjects were who constituted their sample. The Peterson and Barney data showing the differences between adult males and children are illustrated in an F1/F2 plot shown in Figure 1.

Figure 1. Frequency of second formant versus frequency of first formant for vowels spoken by men and children, which were classified unanimously by all listeners (Peterson and Barney, 1952).

Our interest in the development of vowels then developed into three principle questions: the first was whether it was possible to accurately trace the development of vowel sounds from around eighteen months to their adult values; in this we were superceded by the excellent work of Eguchi and Hirsh (1969) whose formant measures for vowels over time are shown in Figure 2.

Figure 2. Mean formant frequencies for combined age groups as shown in the key. Each point represents the combination of Formant 1 and Formant 2 for each of the six vowels. The different symbols together with the lines that join them represent the different ages. The broken circles are drawn around all points for a given vowel. (Eguchi and Hirsh, 1969).

There were, however, subsidiary questions relating to the problem of the ontogeny of vowels, in particular, whether children of the same chronological age but at different stages of physiological development, would demonstrate differences in vowel formant frequencies because of their differences in growth. I will report this information later.

The second principle question related to formant measures of vowels produced by groups of children who were measureably different in their linguistic development, since a great deal of space in the phonological literature has been (and continues to be) devoted to a discussion of how and in what sequence consonant sounds emerge. We felt that children at different stages in the acquisition process might give us some information relating to this question, at least for vowels.
A third, and last principle question, concerned the manner in which vowel sounds are perceived by children when these sounds are produced both by themselves and adults. Since vowels are perhaps more easily acoustically measured than consonants in the output of children, and since there appears to be more listener agreement on their character, we considered this line of investigation was one worth following.

**Studies of Vowels**

Because of their clear separation in the vowel quadrilateral, our energies were directed chiefly to an examination of four vowels: \(/i/\) as in "heed", \(/æ/\) as in "had", \(/ə/\) as in "had", and \(/u/\) as in "who'd", produced by both children and adults, usually in an h—d environment. Other studies reported in the literature have examined a wider set than this. The choice of these vowels, however, allowed us to compare our results with results from numerous studies conducted with adults, and in retrospect, to consider some issues, e.g. individual variation, as they apply to the emergence of vowels during acquisition. Bearing in mind the problems of holding "mechanical" (i.e. child) parameter constant, and the difficulties of minimizing measurement variation (see Kent, 1976, 1978 for details on acoustic analyses of children's vowels), we hoped to view the vowel system "settling down" across chronological age.

In an early paper, Okamura (1966) measured five vowels spoken by 475 Japanese children and demonstrated that the formant frequency construction of these vowels was quite different between children and adults. A copy of his centre formant frequency measurements is shown in Figure 3.

It will be seen that for all of these vowel sounds, the formant frequency measurements appear to plateau around seven years of age. When we came to compare our own data for four-year-old English speaking children with that of Okamura, we found a fair measure of agreement for formant two. Our measurements are shown in Table 1.

Interestingly, the use of duration of vowels in emerging phonology appears, at least in one report (Di Simoni, 1974), to follow this development trend; by age six, durational differences between vowels becoming stabilized in children's speech. This issue is, however, confused and the reader is referred to a comprehensive account of factors in Greenlee (1978).

As mentioned earlier, one of our interests was to determine whether differences in physiological age (whilst holding chronological age constant) would, in fact, change the acoustic characteristics of children's vowel sounds. It did not appear sensible to group children by CA for the purposes of examining vowel development if, in fact, their physiological ages were markedly dissimilar. The motivation for this observation was the assumption that a difference in physiological age would mean a difference in vocal...
tract length and therefore a difference in characteristic vocal tract resonances.

Table 1. Vowel productions: means and standard deviations, in hertz for F1 and F2 measurements of control and experimental groups (Gilbert, 1970).

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Control F1 Mean</th>
<th>Control F1 Stand. dev.</th>
<th>Experimental F1 Mean</th>
<th>Experimental F1 Stand. dev.</th>
<th>Control F2 Mean</th>
<th>Control F2 Stand. dev.</th>
<th>Experimental F2 Mean</th>
<th>Experimental F2 Stand. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td>442</td>
<td>107</td>
<td>2510</td>
<td>99</td>
<td>555</td>
<td>149</td>
<td>2613</td>
<td>67</td>
</tr>
<tr>
<td>/æ/</td>
<td>917</td>
<td>183</td>
<td>1710</td>
<td>251</td>
<td>859</td>
<td>130</td>
<td>1631</td>
<td>122</td>
</tr>
<tr>
<td>/a/</td>
<td>693</td>
<td>112</td>
<td>1246</td>
<td>157</td>
<td>727</td>
<td>113</td>
<td>1216</td>
<td>299</td>
</tr>
<tr>
<td>/u/</td>
<td>539</td>
<td>166</td>
<td>1255</td>
<td>202</td>
<td>533</td>
<td>115</td>
<td>1336</td>
<td>207</td>
</tr>
</tbody>
</table>

We found that both F1 and F2 naturally show a tendency to drop with an increase in chronological age from fourteen to eightyfour months, and that when subjects were reassigned to groups by Bone Age (BA) (Harrison, et al. 1964) groupings, the same pattern emerged. We found no statistical difference between CA and BA on our formant measures; we thus concluded that grouping children by a measure of physiological maturity (rather than CA) does not in the final analysis alter results.

In retrospect, I am not sure that this was an appropriate conclusion to draw, based on the way in which we assigned children to BA groupings. I suspect that it would have been more appropriate to have taken both BA plus skeletal size, i.e. height and weight, and then compared them with children of similar CA and intelligence. We know from the work of Negus (1949) that the larynx develops most rapidly between 3:0 and 5:0 CA and then increases in size to maturity very slowly. This point should have been taken into account. I am still not convinced that we have solved the physiological age problem in our deliberations.

From a consideration of physiological age we then moved to a slightly different view of the process, that is, would children at different levels of linguistic development, but at the same CA exhibit any significant differences in vowel production. Given that children are the same height and weight our assumption would be one of no difference, since we would expect that whatever emerged from the vocal tract would be of the same order, regardless of whether or not each child's linguistic abilities were different.

We recorded children at 4:0 CA divided into groups on the basis of normal and late language usage. Although there were no differences between these groups in terms of mean formant two measurements, when we played tokens produced by the late language users to adult listeners for identification, the adults were definitely confused in their perception, a result which we had certainly not anticipated.

We interpreted this discrepancy as an indication that children who are at a less mature stage of linguistic development are doing "something" to the vowels which cannot be accounted for on an acoustic basis. A thorough examination of the acoustical similarities and dissimilarities between normally developing children and language delayed children is necessary before we can make any further judgements. It may well be that the dynamic acoustic information distributed over the temporal course of the syllable, is affecting listener judgement differently in each case.

The last question to which we addressed ourselves involved the perception of vowels. In 1967, Menyuk reported an experiment in which she showed that the phoneme boundaries for a set of vowels in consonantal context were the same for six children between 5:0 and 10:0 as they were for adult listeners. We found in our experiments that children at 4:0 have no difficulty in discriminating four broadly spaced vowel tokens spoken by themselves and by adults, when these are presented to them in an h-d context. We also found that, when children at this age are asked to produce vowel sounds in an h-d context (in response to the same vowels in h-d context spoken by adults and child speakers other than themselves), there is virtually no difference between F0 and F2 in their tokens, and the tokens of the speakers whom they are imitating.

Lieberman (1978) and his associates at Brown University have data which shows a gradual and consistent improvement in the children's productions of vowels of English from the early stages of babbling through to 3:0; an age at which the children are using meaningful sentences and conversing with the experimenters. Lieberman's data is very robust, and certainly corroborates our own notions about vowel development.
Conclusion

The question of the acoustical development of vowel sounds appears to be reasonably well answered by now. That is, one sees an increasing trend over the first six years towards the adult form in terms of fundamental frequency, F2 and F3. The plateauing between 6:0 and 8:0 is undoubtedly related to the fact that the vocal tract at this time is approaching its adult measurement.

The question of the child's perception of vowel sounds appears more equivocal. Since perception will have to be accounted for by correct usage in production, we will need further experiments of the kind recently reported by Greenlee (1978) before any definitive statements can be made. The same reservation is also true for adult listeners' perceptions of child talkers. There appears to be minimal evidence that children attempt to mimic the acoustic characteristics of the adult speech that they hear, although we do know from Garnica (1974) that at least the mother is adjusting the acoustic characteristics of her utterances to the child. Why is it then that children's vowel utterances are so clearly delineable at a relatively early age? As discussed by Verbrugge et al. (1976) normalization does not appear to be a satisfactory answer.

References


