Published in final edited form as:

J Speech Lang Hear Res. 2014 February 1; 57(1): 1–15. doi:10.1044/1092-4388(2013/12-0112).

A longitudinal study of very young children's vowel production

Rebecca W. McGowan, Richard S. McGowan, and Margaret Denny

CReSS LLC Lexington, MA

Susan Nittrouer

Department of Otolaryngology, The Ohio State University Columbus, OH

Abstract

Purpose—Ecologically realistic, spontaneous adult-directed longitudinal speech data of young children was described by acoustic analyses.

Method—The first two formant frequencies of vowels produced by six children from different American English dialect regions were analyzed from ages 18 to 48 months. The vowels were from largely conversational contexts and were classified according to dictionary pronunciation.

Results—Within-subject formant frequency variability remained relatively constant for the span of ages studied here. It was often difficult to detect overall decreases in the first two formant frequencies between the ages of 30 and 48 months. A study of the movement of the corner vowels with respect to the vowel centroid showed that the shape of the vowel space remained qualitatively constant from 30 through 48 months.

Conclusions—The shape of the vowel space is established early in life. Some aspects of regional dialect were observed in some of the subjects at 42 months of age. The present paper adds to the existing data on the development of vowel spaces by describing ecologically realistic speech.

Introduction

Speech acquisition by children involves all aspects of development including motor control, perception, cognition, and physical growth of body structures. As children acquire the ability to speak in a language environment, they often speak to be understood by adults who possess a working knowledge of that language. The degree to which children who are just learning to speak have an understanding of their language's structure, including its phonology, has not been established. Thus, an important avenue of research is to describe the speech production of children who are learning to speak their ambient language concurrent with gaining a working knowledge of its structure. One way to describe the process of learning vowel production is to measure the acoustic parameters of vowels as they are produced in meaningful utterances directed toward adults.

The present study is a descriptive one, based on acoustic speech data collected for other purposes (Nittrouer, 2010). The first two formant frequencies, F1 and F2, are used to follow the acoustic phonetic development of the monophthongal vowel production of six American children from age 18 to 48 months as they learn to speak English. The speech samples were recorded at six-month intervals in the company of English speaking adults. In the present

study, the vowels are classified according to their dictionary phonemic identity in the words spoken by the children, and, therefore, the so-called standard American English pronunciations are identified. In this way, the acoustic phonetic aspect during the acquisition of the vowel portion of the children's phonological system can be described and documented.

Ecologically realistic, spontaneous adult-directed speech, from both conversational speech, speech elicited by pictures, and a small number of words elicited by imitation with vowels in various phonetic contexts are included in the present study. Formant measurement is performed on words that could be understood by the authors in the context in which they were spoken. The following review of the literature shows that the kind of data examined in the present study is unique. The methods by which the data are collected will be the focus, and the results of many of these studies will be noted later as the results of the present study are discussed.

Several studies have measured vowel formant frequencies from speech produced by children and adults to understand speech development (see Vorperian and Kent [2007] for a review). Peterson and Barney (1952) measured formant frequencies of vowels spoken by children. The ages and place-of-residence of the children were not specified, but they must have been old enough to read from the elicitation lists of /hVd/ words. The classic study of Eguchi and Hirsh (1969) included pre-literate and literate children and measured F1 and F2 of vowels in two short sentences spoken by children ages three to thirteen years and adults from the St. Louis area. Each vowel of interest was contained once in either of these sentences. For children younger than seven years old, the sentences were elicited by imitation; older children and adults read the sentences aloud.

As formant tracking technology has improved, it has been possible to consider large groups of children and to track the trajectories of their vowel formants automatically. Hillenbrand, Getty, Clark, and Wheeler (1995) extended Peterson and Barney's 1952 study of vowels by analyzing the speech of 46 ten- to twelve-year-old children and 93 adults primarily from Michigan. Speakers read lists of /hVd/ words. Looking more broadly at development, Lee, Potamianos, and Narayanan (1999) studied changes in formant frequencies, among other acoustic measures, in the speech of 436 children aged five to eighteen years and adults. Primarily Missouri and Illinois speakers were recorded speaking target /bVt/ words in a carrier sentence as well as five meaningful sentences. The speech tokens were elicited with written words and sentences, except in the case of some five to seven year-olds, who required elicitation by imitation. Formant measurements for children at ages three, five, and seven years and adults from Dallas, Texas were made by Assmann and Katz (2000) for a perceptual study. Due to their study's focus, a subset of the vowels was analyzed based on "adequate pronunciation quality." The vowels were spoken in /hVd/ contexts in imitation of a recording of an adult speaker. As part of a larger study Ménard, Schwartz, Boë, and Aubin (2007) recorded five each of four-year-olds, eight-year-olds, and adults producing Canadian French vowels elicited in a carrier phrase after an adult speaker produced the phrase.

Common to all of the studies reviewed above is the fact that word elicitation was either achieved through reading or by imitation, depending on the age of the subject. Some studies

employed only reading, while others required a mix of these elicitation methods. Some studies have used only elicitation by imitation to study the vowel production of pre-literate four-year-old children (e.g. Kent & Forner, 1979). The phonetic contexts of the vowels of interest were fixed in each of these studies.

Obtaining and processing speech samples of children younger than three years of age is a major difficulty. High and highly variable fundamental frequencies make automatic formant tracking unreliable. For very young children, say under 15 months of age, it is common to study acoustics of vowel-like spontaneous productions without reference to category, as Kent and others have done (e.g. Kent & Murray, 1982; Kent, Osberger, Netsell, & Hustedde, 1987; McGowan, Nitttrouer, & Chenausky, 2008). In studies of slightly older children, categories based on features have been employed for studying spontaneous speech. For example, Gilbert, Robb, and Chen (1997) measured formant frequencies in children over the course of ages 15 to 36 months and vowels were categorized by tongue height and advancement.

The present longitudinal study employs a methodology that differs from previous studies. Six children from 18 months of age to 48 months of age are followed longitudinally in their speech production, and the present study is unique in using largely spontaneous speech where the vowels are grouped according to dictionary phonemic classification. There are various ways to consider the data descriptively. With dictionary classification, the growth of the shape of individual vowel spaces, including aspects of dialect, can be documented as children accumulate their productive vocabularies. Issues of variability are always important when studying the development of children's speech. Comparisons with previous studies that employed other elicitation methods will be made. While there are advantages to the methodology in the present study, this methodology will be unable to follow development of fine phonetic detail, because of the many factors that cannot be controlled when spontaneous speech is the main source of data. The previous studies are advantageous in this respect.

The present paper will explore topics in a descriptive manner. We will consider how the results differ from those of many of the previous studies described above. The studies will be compared in terms of variability and changes of mean formant frequencies with age. It is expected that the vowels in the present study will exhibit greater variability than those of previous studies because spontaneous speech data are employed here. Indications of the development of regional dialect are also examined in the present work. The question as to when the basic shape of the vowel space is established is important and will also be considered. The present paper is unable to explore such effects as gender on vowel production because of the relatively small subject pool (Perry, Ohde, & Ashmead, 2001).

Methods

Participants

The six children studied in the present work were American English speakers from three dialect areas: Northern (one from Rochester, MN; two from Chicago, IL), Midland (one from Oklahoma City, OK), and West (two from Logan, UT). The recordings were originally made in a study investigating language development in children with and without hearing

loss (Nitttrouer, 2010). The speakers reported on here all passed universal newborn hearing screenings, and at 36 months of age passed hearing screenings consisting of the pure tones 0.5, 1.0, 2.0, and 4.0 kHz presented at 20 dB HL to each ear separately.

Material

The children were recorded at six month intervals from 18 to 48 months playing with toys with a parent (spontaneous conversational speech) as well as naming vocabulary items from pictures (expressive vocabulary samples), and repeating words after an experimenter (isolated target words). The latter types of recordings were made only at ages 36, 42, and 48 months as the children's speech intelligibility was tested using the Children's Speech Intelligibility Measure (CSIM) of Wilcox and Morris (1999). The play recordings were made at all ages. The play recordings were approximately 20 minutes long. A fourth type of recording was made of the mother giving direction to the child; the child may often be heard clearly as well in these recordings. These recordings were approximately 10 minutes long and supplemented the above material. The percentage of vowels in this study that are from conversational speech are for each subject, in increasing order, 73%, 73%, 80%, 88%, 93%, and 94%.

Each subject's speech was transcribed orthographically. We were interested in learning about how children produced vowels that were directed toward adults. Vowels found in words that we understood were categorized according to what is referred to in the present work as the *dictionary vowel*, rather than by their pronunciation. Dictionary vowels were determined by the standard pronunciation given in the *American Heritage College Dictionary, 3rd Ed.* Stressed vowels were easily categorized, unless the dictionary gave multiple accepted pronunciations, in which case we did not analyze the vowel. We typically did not measure vowels produced without stress, unless the dictionary pronunciation was /ʌ/ (i.e. we did not measure vowels the dictionary categorized as /ə/ or that the dictionary said were non-reduced vowels if the production was highly reduced).

This approach to categorizing target vowels does not take the dialect's target vowel into account. Thus, when formant frequencies are used to characterize these orthographically transcribed vowels, the effects of dialect can become apparent. For example, perhaps a dictionary /u/ is more fronted for the western speakers than standard pronunciation would dictate and production of /u/ in these speakers may not relate to their ability to produce a backed vowel.

Vowels found in a nasal context (e.g. “mom”) were not analyzed due to the potential for nasal formants; neither were vowels that were intended to be followed by /ɹ/ (e.g. “our”) due to the possibility of r-coloring of the vowel. Stressed vowels following /ɹ/ were measured. If a vowel had a very high F0 (with F0 was higher than the expected F1 frequency region), it was not analyzed due to the difficulty in making accurate formant measures.

Measurements

The sound files, recorded at 48 kHz, were decimated to a sampling rate of 16 kHz. F1 and F2 were measured by hand from the cross-sectional spectrum while simultaneously viewing the spectrogram using the speech analysis software Speech Station II and Wave Surfer. (Two programs were used because operating system changes made the former program unusable.) Spectra were plotted from a DFT analysis with a Hamming window size of 512 samples without preemphasis.

This paper presents only the results from the vowels /i, ɪ, e, æ, ʌ, ɑ, ɔ, ʊ, u/ that were produced as monophthongs. The vowels /e/ and /o/ were not included because they are typically diphthongized in American English. Vowels with steady formants in the middle of the vowel were measured once in the middle 50% of the vowel. A formant frequency was estimated by interpolating by eye among the three highest voice harmonics in the vicinity of the formant in the spectral cross-section. The frequency vicinities of F1 and F2 were estimated from the spectrogram. We also attempted to measure F3 in order to ensure that the measured F2 was not confused with F3. The F3 measurements were not used in the present study, because they could not always be reliably measured. If the vowel was diphthongized with formant movement visible on the spectrogram, it was not analyzed in the present study. An exception to the method described above was made for /u/, which was often highly yodded in some contexts such as /d/. When the formants decreased markedly over the course of /u/, just one measurement was made toward the end of the vowel.

Measurements on each subject were made by one author, and another author made measurements on 10% of these vowels to verify the reliability of the F1 and F2 measurements. The resulting reliability measures are shown in Table I in terms of the percent absolute difference in the two measurements. The average percent difference in F1 (from 6.9% to 9.7%) is about twice that for F2 (3.6% to 4.6%). These reliability measures were deemed acceptable by the authors.

Results

Frequency of vowels

The frequency of vowels measured is not representative of the frequency of production by the subjects, because the analyzed vowels are those within the words that were understood and not excluded for other reasons (nasal context, rhotacization). In general it is true that there are more tokens at the later ages due to a combination of the children speaking more and being more intelligible. The number of tokens for each vowel is shown in Appendix A.

In addition to determining the dictionary vowel, which is used as the basis of analysis in this paper, each vowel was also transcribed phonetically. Actual production is compared to dictionary vowel in Table II. This will be discussed along with observations on dialect features below. The two transcribers were originally from suburban Boston and central New Jersey, both living in the Boston area. Recall that vowels for which the production was reduced or diphthongized are not included in this analysis. The dictionary vowel was different from the transcribed vowel for 8.0 % of dictionary /æ/ tokens, 9.1% of dictionary /i/ tokens, 9.9% of dictionary /ʌ/ tokens, 11.5% of dictionary /u/ tokens, 11.8% of

dictionary /ɪ/ tokens, 14.3% of dictionary /ʊ/ tokens, 14.6% of dictionary /ɑ/ tokens, 14.7% of dictionary /ɛ/ tokens, and 40.6% of /ɔ/ tokens. In terms of transcription, the vowels at the corners of the vowel space were produced more accurately than were the others. With the exception of /ɛ/, the front vowels were produced more accurately than the back vowels. The low back vowels were among the least accurately produced vowels.

Formant frequencies

Vowel spaces at each age are plotted in Figure 1. Mean formant frequencies at each age are given in Table III. Each value shown is the grand mean of the subjects' means of all tokens. The standard deviation, shown in parentheses, is the dispersion of the subjects' means about the grand mean, i.e. the between-subject variability.

Direct comparison is possible between the formant frequency values measured in the present study and those from Eguchi and Hirsh (1969), who measured the formant frequencies in the vowels /i, e, æ, ɔ, u/ of five children at ages 36 and 48 months. Assmann and Katz (2000) measured formant frequencies in vowels of three children at age 36 months. The results from these two studies are plotted with the current results in Figure 2. Bars indicate the standard deviation between subjects in each study.

For mid and low vowels, F1 values are generally lower in Eguchi and Hirsh's data than in the present data. Most notable is a low F1 in /æ/ and the resulting similarity between /æ/ and /ɛ/ in Eguchi and Hirsh's data as compared to the present study at both 36 and 48 months. Otherwise, F1 and F2 values are generally quite similar between the present study and Eguchi and Hirsh's data, as well as Assmann and Katz' data.

Formant frequencies over time

To examine the relation of formant frequencies to time, hierarchical multiple regressions with dependent variables F1 or F2 and factors vowel and then age were computed for each subject using SPSS. The first factor, vowel, was expected to account for a large amount of variation in F1 and F2 because the vowels are distinct in F1–F2 space; adding age determines how much additional variation is accounted for by the subject's age. Hierarchical regression can be effective when the first group of factors, denoted as control variables, is known to have an effect on the dependent variable, and the subsequent group of factors, denoted predictor variables, is the set for which interest is focused. Here the vowel identity is the control variable, and subject age is the predictor variable. Thus, for each subject and formant frequency, the first regression was a one-way ANOVA with vowel identity as a factor. The second regression added age as a factor, which is similar to an ANCOVA, but instead of partialing out the effect of the covariate age, the factor of vowel identity is partialled out to find the effect of age on formant frequency (Bickel & Doksum, 1977; Cohen, Cohen, West, & Aiken, 2003; Field, 2009).

The vowel categories were coded with unweighted effects coding. This coding makes each vowel independent variable a comparison between the vowel and the centroid (the unweighted mean of vowels); specifically the difference between the mean formant frequency for the vowel and the centroid for that age were regressed against age (Cohen et al., 2003). For reasons discussed below in the section on “Shapes and positions of vowel

spaces”, the centroid was determined to be useful for ages 30 months and older in the present data; thus the regressions were computed for ages 30 through 48 months.

The regression analyses were employed as descriptive statistics for the present data set and should not be viewed as predictive statistical models (Hays, 1981 pp. 459–61). Vowel alone accounts for a substantial amount of F1 variation for each subject (between 60.7% and 69.8%, $p < 0.001$). Further, age accounts for a certain amount of additional variation of F1 in three of the subjects (between 0.9% and 5.9% additional variation, $p < 0.001$). These three subjects with statistically significant F1 decrease with age were Chicago, IL 1 (slope = -3.57 Hz/mo), Chicago, IL 2 (slope = -4.57 Hz/ mo), Rochester, MN (slope = -9.16 Hz/ mo). The effects of the addition of age to the model are moderate in two further subjects Logan, UT 2 ($p = 0.003$ and slope = -2.18 Hz/mo) and Logan, UT 1 ($p = 0.005$ and slope = -3.45 Hz/mo).

Vowel accounts for a large amount of the variation in F2 for each subject (between 73% and 81.7%, $p < 0.001$). The addition of age into the model accounts for a small amount of additional variation in two of the subjects (between 0.6% and 0.7% additional variation, $p < 0.001$) These subjects were Chicago I (slope = -8.04 Hz/mo) and Logan II (slope = -9.23 Hz/mo).

Within-subject formant frequency variability

Studies of children's vowel formant frequency development have often concluded that children achieve vowel targets more accurately as they get older because vowel formant frequencies become less variable with age (Eguchi & Hirsh, 1969; Lee et al., 1999; Assmann & Katz, 2000; Ménard et al., 2007). A decrease in acoustic variability is often interpreted as an increase in articulatory precision (Vorperian & Kent, 2007) resulting from the maturation of the motor control system (Smith, 2010). There is also evidence that children's speech is more strongly influenced by intrasyllabic coarticulatory effects than is the speech of adults (Nitttrouer, Studdert-Kennedy, & McGowan, 1989; Nitttrouer, Studdert-Kennedy, & Neely, 1996), which could also influence the measured variability of vowel formant frequencies. In the present study, variability over time in vowels is examined among vowels found in the varied phonetic contexts of running, natural speech.

A measure of F1 or F2 within-subject variability that summarizes all subjects was calculated following Eguchi and Hirsh (1969). For each vowel, the square root of the mean of individual subjects' variances was divided by the grand mean frequency across subjects for that vowel. This is known as *normalized within-subject variability* in the present study and represents typical variability. Results are shown in Figure 3, in which the variability in each vowel is plotted over time.

Comparison of the present results to previous studies in which vowels were produced in constant phonetic contexts may highlight important differences between the types of data. Normalized within-subject variability was also calculated by Eguchi and Hirsh (1969) for subjects beginning at age 3 years, as well as Lee et al. (1999), beginning at age 5 years. The magnitude of normalized within-subject variability is higher in all vowels in the present

study than in Eguchi and Hirsh's data. Normalized within-subject variability in the present 4-year-old data was higher than in the 5-year-old children in Lee et al.

Dialect and phonetic transcriptions

Examination of the vowel space reveals qualitative features among the subjects that could be indicative of dialect region, particularly for the older ages. Figure 4 shows the vowel space of each subject at age 42 months. Dispersion ellipses centered at the mean of each vowel are plotted; the length of the axes of these ellipses is 1.5 standard deviations in either direction from the mean. Ellipses are labeled at their centers.

Some relative placements of the vowel ellipses could be related to dialect. If we apply the criterion that the ellipse for the vowel /æ/ intersects that for /ɪ/ in order to judge whether the /æ/ is raised and fronted, then there is a high degree of raising and fronting of /æ/ for Rochester, MN (Figure 4a). Further, there is a slight degree of raising and fronting of /æ/ for Chicago, IL 1 and Chicago, IL 2 (Figures 4b and 4c) compared to the three non-northern speakers (Figures 4d–f). Rochester, MN and Chicago, IL 2 (Figures 4a and 4c) also have more fronted /ɑ/ than the other speakers. The subject in 4d (Oklahoma City, OK) shows an /ɑ/ - /ɔ/ merger. In general, /ɪ/ and /e/ overlap substantially in all subjects except in Figure 4d (Oklahoma City, OK). The vowel /u/ is more fronted for Oklahoma City, OK and Logan, UT 2 (Figures 4d and 4f) than for Rochester, MN, Chicago, IL 1, and Chicago, IL 2 (Figures 4a–c).

Comparisons between the phonetic transcriptions of the vowels made by the investigators and the dictionary pronunciation are shown in Table II. Given the overlap between /ɪ/ and /e/ in the ellipses in Figure 4, it is not surprising that many dictionary /ɪ/s were transcribed as /e/. Also, there was also a very substantial percentage, 30%, of /ɔ/s transcribed as /ɑ/. At a lesser level of substitutions or confusions were /ɪ/ for /i/ at 8%, /ʊ/ for /u/ at 6%, and /ɑ/ for /ʌ/ at 4%.

Shapes and positions of vowel spaces

We are interested in comparing the development of each child's vowel space relative to a feature that depends on all the data defining the individual's vowel space. Chung et al. (2012), in a study of five languages, found that language-specific corner vowels are well-established in relation to the centroid by the age of five years. The centroid is the unweighted mean of vowel locations in log-transformed F1–F2 space. Here we consider the shapes of individual vowel spaces as in Chung et al. (2012).

Because the children in the present study did not always produce tokens of every vowel, the location of the centroid was calculated only when most vowels were present, 30 months and older. Figure 5 shows the corner vowels and centroids from ages 30 months and older superimposed. Age is indicated by symbol shape. The centroid decreases slightly in both F1 and F2 over time for most subjects. Variability in /u/ stands out in several subjects as being larger than in the other corner vowels. The /ɑ/ vowel of Logan, UT 1 (Figure 5e) has a substantial trajectory through time.

To better understand the stability or variation in the relation of the vowels to the centroid, variability from age 30 to 48 months in the direction from the centroid to the vowel was calculated. The standard deviation of direction was calculated and normalized by multiplying by the mean distance because variability in angle is magnified by distance to the centroid. Results are shown in Figure 6. The back corner vowels, /u/ and /ɑ/, are more variable in direction for some subjects than are the front corner vowels, /i/ and /æ/. The back vowel /u/ is highly variable for some subjects.

Discussion

This section highlights the findings for vowel production when the tokens extracted are largely from recordings taken in conversational settings. Comparisons are made with previous studies of vowel production that employed different kinds of elicitation methods. These methods, as reviewed in the Introduction, did not include natural conversational situations. We begin with a discussion of within-subject variability, because this highlights differences between the current data and the previous data.

Within-subject formant frequency variability

Comparison of normalized within-subject variability to previous works highlights important differences in the type of data analyzed. Overall levels of normalized within-subject variability in F1 and F2 are found to be much higher in the present study than in Eguchi and Hirsh (1969, ages 3 and 4 years) or in Lee et al. (1999, age 5 years). Further, while previous studies have found that normalized within-subject variability decreases with age (Eguchi & Hirsh, 1969; Lee et al., 1999), no apparent decrease in F1 variability, nor in F2 variability, is detected over time in the present study (Figure 3). On the other hand, for each vowel except /u/, overall levels of normalized within-subject variability are lower in F2 than in F1 in the present study, as well as in Eguchi and Hirsh (1969) and Lee et al. (1999).

The greater amount of pronunciation variability found for this study compared to those earlier ones can likely be attributed to several factors, including variability in segmental phonetic contexts, which would provide greater opportunities for coarticulation; supra-segmental factors arising from speech being spoken in natural contexts; emotional state; and pragmatic factors associated with these meaningful productions, as opposed to the elicited forms used in Eguchi and Hirsh (1969) and in Lee et al. (1999). While articulatory precision increases with age, as noted by Eguchi and Hirsh, it appears to be overcome in the present work by the factors just noted as the children's productive vocabulary grows.

The idea that some of the variability can be explained by coarticulation of adjacent phonetic segments is supported by findings from studies with adults (Stevens & House, 1963; Hillenbrand, Clark, & Nearey, 2001). To determine the extent to which coarticulation causes an increase in within-subject variability, a baseline variability –normalized within-subject variability for all vowels following an alveolar consonant– was calculated. This is shown in Table IV. The percentage of normalized within-subject variability for utterances in alveolar context compared to the within-subject variability in all contexts is less than 85% in F1 for vowels /ʌ/ and /ɑ/, and in F2 for vowels /i/, /ɪ/, /e/, /æ/, /ʌ/, /ɑ/, and /ɔ/. Only the back vowels /ɔ/ and /u/ do not appear to exhibit any additional variability in contexts other than

initial alveolar consonant context. In general the results of Stevens and House (1963) indicate that alveolar context perturbs high vowels' F1 more than low vowels' in relation to the other places of articulation, and that is consistent with the results shown in Table IV. Further, it is known that change in F2 from alveolar context is largest for the vowels /*o*/ and /*u*/ for adults (Stevens & House, 1963; Hillenbrand et al. 2001). If children behave similarly to adults in this respect, then the variation in F2 in other consonant contexts is smaller than in the initial alveolar context for these vowels. We conclude that coarticulatory effects, interacting with other factors, do make a substantial contribution to the intra-subject variability in the present data.

Other factors that interact with phonetic context that need to be recognized to contribute to formant variability are supra-segmental and pragmatic factors. For example, the hyper- and hypo speech theory of Lindblom (1990), which involves the notion of sufficient discriminability and changes to speech behavior due to pragmatic factors, should have direct bearing on the variability seen in these children who are either interacting with a caregiver, an examiner, or talking to themselves. Further, in a conversational context, these young children exhibited a variety of emotional states that would certainly influence the acoustics of their vowels.

Formant frequencies

Formant frequencies measured in the present study are qualitatively similar to those measured at ages 36 and 48 months by Eguchi and Hirsh (1969) and at 36 months by Assmann and Katz (2000), as shown in Figure 2. In particular, F2 measurements are quite similar. Qualitative differences are found between the present F1 measurements and those reported by Eguchi and Hirsh in the mid and low vowels, and in particular a large difference in the F1 of /*æ*/ is found between the two studies. Previous studies (e.g. Hillenbrand et al., 1995, comparing formant frequencies from ages 10–12 years to Peterson and Barney, 1952) have found differences among data sets that have been attributed to differences in methodology, dialect differences, or the passage of time. In children's speech, F1 is difficult to measure accurately due to high fundamental frequencies and the resulting wide spacing of harmonics. The validation of the formant measurements in the present study shows that differences in measurements made independently are about twice as large for F1 than for F2 on a percentage basis. This may be one reason for the differences in F1 reported here and in Eguchi and Hirsh.

In addition, speakers' dialects may be a factor in the values of the measured formants. The largest difference between the present data and Eguchi and Hirsh is the similarity of /*æ*/ to /*ε*/ in the latter. The speakers in that study were from the St. Louis, MO area, which Labov, Ash, and Boberg (2006) have found is beginning to show the Northern Cities Chain Shift. The majority of speakers in Labov et al.'s study were in their 20s to 40s when the study was conducted in the 1990s, putting them in the same age group as the children in Eguchi and Hirsh's (1969) study. The similarity between /*æ*/ and /*ε*/ is not seen in the speakers from Texas in Assmann and Katz, or in the average across speakers from different dialect areas in the present study.

Formant frequencies over time

Hierarchical multiple regressions on F1 and F2 were calculated for each subject. Vowel is found to account for a large amount of variability in F1 and F2 for all subjects; this simply confirms that vowels differ from one another in F1–F2 space.

The addition of age to the model accounts for substantial additional variability in F1 for three subjects and for additional variability in F2 for two subjects. Each of these subjects shows decreases in the formants, which are expected due to the lengthening of the vocal tract with age. The decreases in F1 are smaller overall than the F2 decreases. However, for the subjects who exhibited notable decreases, the percentage decrease given by the regression slopes are 8%, 10%, and 19% for F1 with an average F1 of about 850 Hz at 30 months, and they are 6% and 7% for F2 with an average F2 of about 2250 Hz at 30 months. This finding is consistent with that of Eguchi and Hirsh (1969), who found that F1 was more stable than F2 over the course of age 3 years to adult. This is likely simply because F1 is smaller in absolute magnitude.

Thus we find some decrease in formants with time, but not for every subject. In particular, given that the F2 decreases are larger in absolute terms, it is surprising that we do not see an effect in more subjects. One factor that could influence the results of the regressions, which are over all vowel tokens, is that the distribution of vowel tokens may shift over time (see Appendix A). Examination of Appendix A indicates that for Rochester, MN, Chicago, IL 1, and Logan, UT 2, there are not large changes in the proportions of various vowels. The subjects from Chicago, IL 1 and Logan, UT 2 are the two subjects showing decreases in F2. The other subjects, however, exhibit an increase in the number of front vowels as compared to back vowels over time. In other words, the mean F2 may have increased somewhat over time for these subjects. This would be one reason that only two subjects, both with little change in the proportion of front to back vowels, show decreases in F2 over time.

Further, the present study examines vowels on a relatively short time scale; consideration of previous data on a similar time scale is revealing. For example, although Assmann and Katz (2000) found an overall statistically significant decrease of formant frequency over the course of age 3 years to adult, there were not always decreases in formant frequencies over the course of the two years between age 3 and 5 years, despite selection of vowels for good pronunciation. F1 values increased for /i, e, a/ and F2 values increased or remained stable for /i, I, e, æ, ʌ, ɔ, u/. Lee et al. (1999) also found some increases in formant frequencies for children between age 5 years and 6 years, despite an overall trend of decreases between age 5 years and adults. On a short time scale, overall decreases in formant frequencies may not be detected without a significantly larger number of measurements.

Interactions of the individual vowels with age were not included in the present analysis because they resulted in high levels of collinearity; regressions for individual vowels were not computed because more tokens would have been necessary. In order to understand the relationship of formant frequencies to age, Assmann and Katz (2000) computed an ANOVA which showed main effects of both vowel and age on the formant frequencies, and which included an age-by-vowel interaction, but this interaction was not significant. In other words, vowels changed over time, but all in a similar manner. Future studies analyzing

larger numbers of tokens should consider whether all vowels are similar in their rates of change.

Dialect

Previous studies of children's vowels have either limited the majority of speakers to one region (e.g. Eguchi & Hirsh, 1969; Hillenbrand et al., 1995; Lee et al., 1999), or have not considered regional differences (e.g. Peterson & Barney, 1952). Comparison of children's vowel spaces from multiple regions does reveal differences that may be due to dialect. Figure 4 shows the vowel space of each subject at 42 months. Three subjects are from the Northern dialect region (Figures 4a–c), one is a Midland speaker, who lives adjacent to a region of Southern speakers, (Figure 4d), and two are from the West (Figures 4e–f).

One of the most noted features of Northern dialects is the Northern Cities Chain Shift, which involves raising and fronting of /æ/, lowering and backing of /e/, lowering and fronting of /ɔ/, fronting of /ɑ/, as well as the backing of /ɪ/, /ʌ/, and /ʊ/ (Labov et al., 2006). We observed in the Results, under “dialect and phonetic transcriptions,” that the three Northern speakers, Rochester, MN, Chicago, IL 1, and Chicago, IL 2 showed at least a small degree of raising and fronting of /æ/. In comparing the F2 of /e/ to F2 of /ɑ/, it appears that /ɑ/ can be considered fronted and/or /e/ backed for speakers Rochester, MN and Chicago, IL 2 (Figures 4a and 4c). The position of /ɔ/ does not appear to be different for these Northern speakers from the other speakers, except that speaker Chicago, IL 1 (Figure 4b) shows more /ɑ/ - /ɔ/ merger than expected for a Northern speaker. Also his u appears to be more fronted than the other two Northern speakers. (This speaker heavily yodded his us after alveolar consonants.) There does not appear to be any extra backing of /ɪ/, /ʌ/, and /ʊ/ for any of the three Northern speakers. All three Northern speakers could be showing what Labov et al. (2006) call the triggering event for the Northern Cities Chain Shift: the raising and fronting of /æ/. Two of these speakers seem to show the closely related /ɑ/ fronting. Hillenbrand et al. (1995) also found a raised and fronted /æ/ in their older Michigan speakers, which they attributed to the Northern Cities Chain Shift.

Labov et al. (2006) classify Oklahoma City, OK as part of the Midland region. This particular area shows a partial /ɑ/ - /ɔ/ merger and fronting of /o/ (we did not study the vowel /o/). The Midland speaker from Oklahoma City, OK (Figure 4d) exhibits the most complete /ɑ/ - /ɔ/ merger of all of the subjects. This speaker also exhibits u fronting, which is a characteristic of Midland dialect. Oklahoma City, OK is adjacent to a region of Southern speakers. This speaker showed the least overlap in /ɪ/ and /e/, so either the speaker is uninfluenced by the Southern dialect or the separation is due to the fact that nasal contexts were excluded.

Western dialect features a merger of the vowels /ɑ/ and /ɔ/ and fronting of u (Labov et al., 2006). Only one Western speaker, Logan, UT 2, exhibits possible u fronting (Figure 4f). There are not enough data to determine whether speaker Logan, UT 1 shows an /ɑ/ - /ɔ/ merger (Figure 4e), and Logan, UT 2 does not appear to show this merger.

In summary, some of the speakers in the present study showed dialect features of the region in which they reside. At the age of 42 months one could expect that dialect would be greatly

influenced by the speech of the caregivers, whose dialect we did not attempt to characterize. This influence may be one reason that some of the children did not exhibit features of their local dialect.

Shape and position of the vowel space

The location of the centroid in log-transformed F1-F2 space was found and plotted in Figure 5. The centroid formants decrease over time somewhat, but they provide a well-defined reference against which to compare the vowels. The corner vowels show some variability in their relation to the centroid but no consistent trends emerge. This suggests that there is no rearrangement of the vowel space in relation to the centroid. The relation of the corner vowels to one another appears to be established early in the development of the vowel space, even as formant frequencies continue to decrease as an expansion of vocabulary and phonetic contexts occur, and as development of suprasegmental and pragmatic competence contributes to a large amount of variability in the realization of the vowels.

The variability of each vowel in the vowel space was considered by quantifying the relation of the vowels to the centroid. Variability in direction from the centroid, shown in Figure 6, reveals some differences in the vowels. The corner vowels are consistently located in relation to the centroid in terms of direction with the exception of one speaker who has a large amount of variability in direction for / α /. Within the set of four corner vowels, the front vowels are the most consistent. The non-corner back vowels tend to be more variable in direction to the centroid. Ménard et al. (2007) measured larger dispersion ellipses (in Bark) for back vowels compared to front vowels for French speaking 4 and 8 year-old children. This difference in area was found to be larger for the children compared to the adults.

Conclusion

This study has measured F1 and F2 of children's vowels in a naturalistic context at the youngest ages at which they are acquiring a vocabulary. The results compare well with previous such studies of older children's speech. The vowel spaces look qualitatively similar to those found for older children in previous work, but the present study also highlights inter-speaker dialect differences not previously examined in children under four years old.

Within-subject variability in young children's vowel production in these mostly conversational environments is higher than that in previous studies. Further, it remained relatively constant through the study, compared with the previous studies with elicited speech, where consistency increased with age. That conversational speech in young children produces more variable within-speaker variability than elicited forms is not surprising. That this variability does not appear to change with age is an important finding. Variability can be caused by coarticulation with surrounding phonetic segments, as well as the important factors of intonation and prosody, such as sentence focus. The consequences of pragmatic factors in variability should also be taken into account. It appears that the increase in the number of phonetic segmental contexts, as well as possible increases in the intonational and prosodic environments in which the vowel segments are embedded counteracts any increase in pronunciation precision with age found in citation form stimuli. An interesting question to

consider is whether, as the previous results suggest, formant frequency within-subject variability in adults is also lower than in children when the variety of contexts found in conversational speech is considered.

There are some substantial decreases in overall F1 and F2 for some subjects over the span of 30 to 48 months. According to Magnetic Resonance Imaging data from Vorperian and colleagues (Vorperian, Kent, Gentry, & Yandell 1999; Vorperian et al. 2005), the vocal tract increases rapidly in length from approximately 7.5 cm (or 47% of adult size) in a newborn to 9.5 cm (or 62% of adult size) at 18 months. The vocal tract then continues to steadily lengthen somewhat more slowly to attain about 68% of adult size at 48 months of age (Vorperian et al. 2005, Fig. 4), which means that there is a 10% increase in average vocal tract length between 18 and 48 months. From 30 to 48 months, the increase is even smaller, at about 5%. Overall, this increase in vocal tract length would mean that, at 48 months, formant frequencies are reduced 95% of their values at 30 months, if vocal tract length were the only factor. For the subjects for whom changes in F1 or F2 were detected, the reduction was always greater than 5%, sometimes by a substantial amount. With such small changes due to vocal tract lengthening, it is not surprising that we are unable to systematically detect the acoustic effect of length changes that occur between the ages of 30 and 48 months for all subjects. These considerations are based on group averages, so there are children who will exhibit even less than 5% vocal tract lengthening. Further, it is not possible from these data to infer the reasons that some subjects do not show decreases in formant frequencies that would be expected with the growth of the vocal tract. Only the fact that the vowel inventory changed for some of the speakers can be offered as a reason, but this does not discount other possible causes, including social influences.

Despite the many changes occurring in children's speech over the course of 30 to 48 months, the relation of the corner vowels to the centroid in the vowel space remains a qualitatively constant feature. Although further changes such as lowering of formant frequencies and improvement in pronunciation precision are expected beyond age 48 months, the overall relations of vowels to one another may be well established at a young age. However, the differences in variability between back and front vowels for these children warrants further research.

Acknowledgments

This work was supported by grants R01 DC001247 to CReSS LLC and R01 DC006237 to Susan Nittrouer.

Appendix

Appendix A

Number of vowels measured for each subject, at each age.

Age (mos)	Subject	i	I	e	æ	Λ	ɑ	ɔ	ʊ	u
18	Rochester, MN	4	0	0	0	0	0	0	0	0
	Chicago, IL 1	2	1	0	1	0	0	0	0	0
	Chicago, IL 2	2	0	0	0	7	0	7	0	1

Age (mos)	Subject	i	I	e	æ	Λ	ɑ	ɔ	ʊ	u
24	Oklahoma City, OK	1	0	0	0	2	0	1	0	0
	Logan, UT 1	4	0	4	1	1	0	0	0	0
	Logan, UT 2	4	0	1	2	0	32	0	0	0
	Total	17	1	5	4	10	32	8	0	1
	Rochester, MN	3	2	4	2	2	0	0	0	2
	Chicago, IL 1	37	12	28	6	26	3	2	2	9
	Chicago, IL 2	9	2	14	6	7	1	0	2	5
	Oklahoma City, OK	20	3	2	14	0	3	0	0	0
	Logan, UT 1	8	2	4	2	0	2	6	0	1
	Logan, UT 2	11	19	1	6	1	0	0	1	0
30	Total	88	40	53	36	36	9	8	5	17
	Rochester, MN	18	6	4	9	3	0	0	2	8
	Chicago, IL 1	68	39	37	25	18	6	9	3	22
	Chicago, IL 2	13	11	5	28	6	4	0	5	15
	Oklahoma City, OK	28	14	11	15	4	6	6	2	6
	Logan, UT 1	11	15	3	9	7	0	0	0	7
	Logan, UT 2	30	49	6	9	6	1	0	0	5
	Total	168	134	66	95	44	17	15	12	63
	Rochester, MN	29	29	9	14	13	3	1	5	12
	Chicago, IL 1	56	76	18	19	28	13	10	13	20
36	Chicago, IL 2	34	55	9	22	21	7	3	8	13
	Oklahoma City, OK	16	21	12	27	4	8	5	3	11
	Logan, UT 1	11	24	9	11	8	2	2	0	10
	Logan, UT 2	37	50	21	30	10	5	0	13	23
	Total	183	255	78	123	84	38	21	42	89
	Rochester, MN	16	29	6	11	6	4	3	4	3
	Chicago, IL 1	57	85	20	25	26	18	5	8	26
	Chicago, IL 2	52	63	11	42	12	24	5	18	32
	Oklahoma City, OK	38	73	6	37	14	9	6	23	10
	Logan, UT 1	27	45	10	13	17	4	2	4	7
42	Logan, UT 2	42	52	18	53	24	9	3	13	37
	Total	232	347	71	181	99	68	24	70	115
	Rochester, MN	36	33	7	19	12	4	4	8	13
	Chicago, IL 1	61	72	23	20	24	8	10	12	20
	Chicago, IL 2	60	61	21	25	20	16	6	10	29
	Oklahoma City, OK	61	50	30	35	28	23	20	9	13
	Logan, UT 1	45	47	11	30	14	6	1	3	20
	Logan, UT 2	56	100	10	48	35	10	11	18	38
	Total	319	363	102	177	133	67	52	60	133
	Rochester, MN	36	33	7	19	12	4	4	8	13
48	Chicago, IL 1	61	72	23	20	24	8	10	12	20
	Chicago, IL 2	60	61	21	25	20	16	6	10	29
	Oklahoma City, OK	61	50	30	35	28	23	20	9	13
	Logan, UT 1	45	47	11	30	14	6	1	3	20
	Logan, UT 2	56	100	10	48	35	10	11	18	38
	Total	319	363	102	177	133	67	52	60	133
	Rochester, MN	36	33	7	19	12	4	4	8	13
	Chicago, IL 1	61	72	23	20	24	8	10	12	20
	Chicago, IL 2	60	61	21	25	20	16	6	10	29
	Oklahoma City, OK	61	50	30	35	28	23	20	9	13
	Logan, UT 1	45	47	11	30	14	6	1	3	20
	Logan, UT 2	56	100	10	48	35	10	11	18	38
	Total	319	363	102	177	133	67	52	60	133

References

- American heritage college dictionary. 3rd ed.. Houghton Mifflin; Boston, MA: 1993.
- Assmann PF, Katz WF. Time-varying spectral change in the vowels of children and adults. *Journal of the Acoustical Society of America*. 2000; 108:1856–1866. [PubMed: 11051512]
- Bickel, PJ.; Doksum, KA. *Mathematical statistics: Basic ideas and selected topics*. Holden-Day, Inc; San Francisco, CA: 1977. p. 248-287.
- Chung H, Kong EJ, Edwards J, Weismer G, Fourakis M, Hwang Y. Cross-linguistic studies of children's and adults' vowel spaces. *Journal of the Acoustical Society of America*. 2012; 131:442–454. [PubMed: 22280606]
- Clopper CG, Pisoni DB, deJong K. Acoustic characteristics of the vowel systems of six regional varieties of American English. *Journal of the Acoustical Society of America*. 2005; 118:1661–1676. [PubMed: 16240825]
- Cohen, J.; Cohen, P.; West, SG.; Aiken, LS. *Applied multiple regression/correlation analysis for the behavioral sciences*. 3rd ed.. Lawrence Erlbaum Associates, Inc; Mahwah, NJ: 2003. p. 158-162.p. 320-327.
- Eguchi S, Hirsh IJ. Development of speech sounds in children. *Acta Otolaryngologica Supplementum*. 1969; 257:307–356.
- Field, A. *Discovering statistics using SPSS*. 3rd ed.. SAGE Publications; London: 2009. p. 212p. 349-351.p. 408-412.
- Fitch WT, Giedd J. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*. 1999; 106:1511–1522. [PubMed: 10489707]
- Gilbert HR, Robb MP, Chen Y. Formant frequency development: 15 to 36 months. *Journal of Voice*. 1997; 11:260–266. [PubMed: 9297669]
- Hays, WL. *Statistics*, third edition. Holt, Rinehart, and Winston; New York: 1981.
- Hillenbrand JM, Clark MJ, Nearey TM. Effects of consonant environment on vowel formant patterns. *Journal of the Acoustical Society of America*. 2001; 109:748–763. [PubMed: 11248979]
- Hillenbrand JM, Getty LA, Clark MJ, Wheeler K. Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*. 1995; 97:3099–3111. [PubMed: 7759650]
- Jacewicz E, Fox RA, Salmons J. Regional dialect variation in the vowel systems of typically developing children. *Journal of Speech, Language, and Hearing Research*. 2011; 54:448–470.
- Kent RD, Forner L. Developmental study of vowel formant frequencies in an imitation task. *Journal of the Acoustical Society of America*. 1979; 65:208–217. [PubMed: 422815]
- Kent RD, Murray AD. Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *Journal of the Acoustical Society of America*. 1982; 72:353–365. [PubMed: 7119278]
- Kent RD, Osberger MJ, Netsell R, Hustedde CG. Phonetic development in identical twins differing in auditory function. *Journal of Speech and Hearing Research*. 1987; 52:64–52.
- Labov, W.; Ash, S.; Boberg, C. *The atlas of North American English: Phonetics, phonology and sound change*. Mouton de Gruyter; New York, NY: 2006. p. 28p. 135p. 60-91.p. 187-197.p. 263-269.p. 279-280.
- Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America*. 1999; 105:1455–1468. [PubMed: 10089598]
- Lindblom, B. Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, WJ.; Marchal, A., editors. *Speech production and speech modelling*. Kluwer Academic Publishers; Dordrecht, The Netherlands: 1990. p. 403-439.
- McGowan RS, Nitttrouer S, Chenausky K. Speech production in 12-month-old children with and without hearing loss. *Journal of Speech, Language, and Hearing Research*. 2008; 51:879–888.
- Ménard L, Schwartz J-L, Boë L-J, Aubin J. Articulatory-acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. *Journal of Phonetics*. 2007; 35:1–19.

- Nittrouer, S. Early development of children with hearing loss. Plural Publishing; San Diego, CA: 2010. p. 1-359.
- Nittrouer S, Studdert-Kennedy M, McGowan RS. The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research*. 1989; 32:120–132. [PubMed: 2704187]
- Nittrouer S, Studdert-Kennedy M, Neely ST. How children learn to organize their speech gestures: Further evidence from fricative-vowel syllables. *Journal of Speech and Hearing Research*. 1996; 39:379–389. [PubMed: 8729924]
- Perry TL, Ohde RN, Ashmead DH. The acoustic bases for gender identification from children's voices. *Journal of the Acoustical Society of America*. 2001; 109:2988–2998. [PubMed: 11425141]
- Peterson GE, Barney HL. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*. 1952; 24:175–184.
- Smith, A. Neural control of orofacial movements in speech. In: Hardcastle, WJ.; Laver, J.; Gibbon, FE., editors. *Handbook of phonetic sciences*. 2nd ed.. Wiley-Blackwell; Chichester, UK: 2010. p. 251-296.
- Stevens KN, House AS. Perturbation of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*. 1963; 6:111–127. [PubMed: 13984031]
- Vorperian HK, Kent RD, Gentry LR, Yandell BS. Magnetic resonance imaging procedures to study the concurrent anatomic development of the vocal tract structures: Preliminary results. *International Journal of Pediatric Otorhinolaryngology*. 1999; 49:197–206. [PubMed: 10519699]
- Vorperian HK, Kent RD. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*. 2007; 50:1510–1545.
- Vorperian HK, Kent RD, Lindstrom MJ, Kalina CM, Gentry LR, Yandell BS. Development of vocal tract length during early childhood: A magnetic resonance imaging study. *Journal of the Acoustical Society of America*. 2005; 117:338–350. [PubMed: 15704426]
- Vorperian HK, Wang S, Chung MK, Schimek EM, Durtschi RB, Kent RD, Ziegert AJ, Gentry LR. Development of the oral and pharyngeal portions of the vocal tract: An imaging study. *Journal of the Acoustical Society of America*. 2009; 125:1666–1678. [PubMed: 19275324]
- Wilcox, K.; Morris, S. *Children's speech intelligibility test*. Psychological Corporation; San Antonio, TX: 1999.

Figure 1.

Vowel spaces from grand mean formant frequencies at (a) 18 months; (b) 24 months; (c) 30 months; (d) 36 months; (e) 42 months; and (f) 48 months. Mean formant frequencies for each vowel are plotted with one standard deviation bars in the F1 and F2 directions.

Figure 2.

Comparison of formant frequency measurements from the present study (a) at age 36 months, with Eguchi and Hirsh (1969) and Assmann and Katz (2000); and (b) at age 48 months, with Eguchi and Hirsh (1969). Mean formant frequencies for each vowel are plotted with one standard deviation bars in the F1 and F2 directions. The data summarize five to six subjects in the present study, five subjects in Eguchi and Hirsh's, and three in Assmann and Katz'. Because Assmann and Katz reported standard error, standard deviation is calculated from those values by $SD = SE \times 3$, where 3 is the sample size from which the mean is estimated. E&H = Eguchi and Hirsh (1969), A&K = Assmann and Katz (2000).

Figure 3.

Normalized within-subject variability for each vowel over time. (a) and (c) F1; (b) and (d) F2. The top two panels show the corner vowels and the bottom two panels show the non-corner vowels for ease of reading. Normalized within-subject variability is the square root of the mean of the individual variances, divided by the grand mean formant frequency.

Figure 4.

Vowel spaces of subjects at 42 months. Dispersion ellipses are centered at the mean of each vowel and axes are 1.5 standard deviations in either direction about the mean in length. (a) Northern: Rochester, MN; (b) Northern: Chicago, IL 1; (c) Northern: Chicago, IL 2; (d) Midland: Oklahoma City, OK; (e) Western: Logan, UT 1; (f) Western: Logan, UT 2. Dialect classifications are from Labov et al. (2006).

Figure 5.

Centroid and corner vowels in log-transformed F1–F2 space, ages 30–48 months. Age is indicated by symbol shape. (a) Rochester, MN; (b) Chicago, IL 1; (c) Chicago, IL 2; (d) Oklahoma City, OK; (e) Logan, UT 1; (f) Logan, UT 2.

Figure 6.
Variability in direction of vowels from the centroid, normalized by distance.

Table I

Reliability of the formant frequency measures as a percent of absolute difference.

Subject	Rochester, MN		Chicago, IL 1		Chicago, IL 2	
Formant	F1	F2	F1	F2	F1	F2
Average	6.9%	4.6%	7.8%	4.0%	7.8%	4.5%
Standard Deviation	5.4%	5.4%	6.6%	4.2%	6.8%	4.4%
N	51	51	158	158	117	117

Subject	Oklahoma City, OK		Logan, UT 1		Logan, UT 2	
Formant	F1	F2	F1	F2	F1	F2
Average	9.7%	4.5%	8.0%	4.5%	8.2%	3.6%
Standard Deviation	9.5%	6.7%	5.2%	3.8%	7.1%	2.9%
N	93	93	68	68	136	136

Table II

Dictionary vowel (across top) compared to vowel observed (along left).

Dictionary		i	I	e	æ	ʌ	ɑ	ɔ	ʊ	u
Observed	i	914	26	3	1	0	0	0	0	3
	I	80	1005	15	6	6	0	0	2	3
	e	2	88	320	16	4	1	0	2	0
	æ	2	2	10	567	0	24	4	0	0
	ʌ	1	8	12	5	366	4	4	15	5
	ɑ	1	3	13	19	17	197	38	1	0
	ɔ	0	0	1	0	1	4	76	1	2
	ʊ	0	6	0	0	10	0	0	162	26
	u	0	1	0	0	2	0	0	4	374
	e	7	1	0	1	0	0	0	0	3
	o	0	0	1	1	0	1	6	2	2

Table III

Formant frequencies at all ages: mean (standard deviation = between-subject ability), n = number of subjects with tokens.

Age (mos)		18	24	30	36	42	48
i	n	6	6	6	6	6	6
	F1	514 (51)	527 (38)	523 (31)	518 (49)	508 (45)	491 (43)
	F2	3418 (197)	3458 (246)	3430 (175)	3386 (265)	3429 (221)	3386 (226)
I	n	1	6	6	6	6	6
	F1	1506 (-)	811 (122)	718 (94)	704 (45)	678 (60)	631 (29)
	F2	2214 (-)	2753 (271)	2735 (183)	2678 (247)	2656 (118)	2648 (279)
e	n	2	6	6	6	6	6
	F1	631 (26)	1038 (143)	854 (106)	898 (83)	819 (71)	767 (40)
	F2	2539 (125)	2610 (230)	2598(284)	2468 (238)	2544 (135)	2513 (174)
æ	n	3	6	6	6	6	6
	F1	1083 (448)	1066 (165)	1032 (88)	1088 (93)	1062 (79)	1015 (128)
	F2	2430 (384)	2638 (370)	2440 (201)	2378 (216)	2431 (155)	2378 (164)
Λ	n	3	4	6	6	6	6
	F1	1057 (85)	954 (155)	940 (125)	890 (44)	904 (81)	841 (52)
	F2	2127 (131)	2142 (203)	1962 (210)	2029 (190)	2071 (124)	1977 (189)
ɑ	n	1	4	5	6	6	6
	F1	1211 (-)	1157 (189)	1264 (196)	1110 (132)	1143 (135)	1049 (106)
	F2	2607 (-)	2020 (278)	1960 (187)	1911 (143)	1848 (276)	1865 (187)
ɔ	n	2	2	2	5	6	6
	F1	986 (49)	899 (200)	927 (192)	898 (117)	914 (149)	959 (138)
	F2	1942 (170)	1474 (136)	1588 (101)	1530 (115)	1620 (206)	1517 (176)
o	n	0	3	4	5	6	6
	F1	-	819 (118)	753 (46)	750 (43)	751 (98)	701 (47)
	F2	-	1976 (309)	2044 (697)	1974 (250)	2050 (138)	1815 (222)
u	n	1	4	6	6	6	6
	F1	515 (-)	661 (122)	597 (47)	560 (34)	525 (45)	512 (42)
	F2	1151 (-)	1534 (149)	1496 (201)	1626 (275)	1446 (266)	1484 (280)

Table IV

Normalized within-subject variability in F1 and F2 for vowels in alveolar context and the percent of normalized within-subject variability in alveolar context compared to all contexts.

	i	I	e	æ	Λ	ɑ	ɔ	ʊ	u
F1 variability	0.142	0.175	0.182	0.150	0.137	0.097	0.120	0.161	0.179
F1 % variability	92	103	91	88	74	69	93	106	106
F2 variability	0.078	0.127	0.128	0.075	0.147	0.101	0.092	0.162	0.236
F2 % variability	50	75	64	44	83	73	72	102	140