Timing of Palatal Surgery and Speech Outcome

Kathy L. Chapman, Ph.D., Mary A. Hardin-Jones, Ph.D., Jeffrey A. Goldstein, M.D., Kelli Ann Halter, M.S., Robert J. Havlik, M.D., Julie Schulte, M.A.

Objective: **To examine the impact of age and lexical status at the time of primary palatal surgery on speech outcome of preschoolers with cleft palate.**

Participants: **Forty children (33 to 42 months) with nonsyndromic cleft palate participated in the study. Twenty children (Group 1) were less lexically advanced and younger (mean age = 11 months) and 20 children (Group 2) were more lexically advanced and older (mean age = 15 months) when palatal surgery was performed.**

Main Outcome Measures: **Samples of the children's spontaneous speech were compared on 11 speech production measures (e.g., size of consonant inventory, total consonants correct, % correct for manner of articulation categories, compensatory articulation usage, etc.). Next, listeners rated a 30-second sample of each child's connected speech for articulation proficiency and hypernasality, separately, using direct magnitude estimation (DME).**

Results: **Group differences were noted for 4 of the 11 speech production measures. Children in Group 1 exhibited larger consonant inventories (and true consonant inventories) and more accurate production of nasals and liquids compared to children in Group 2. On the DME task, significant group differences were found for ratings of articulation proficiency and hypernasality. Children in Group 1 exhibited better articulation and less hypernasality than children in Group 2.**

Conclusions: **The findings suggested that children who were less lexically advanced and younger at the time of palatal surgery exhibited better articulation and resonance outcomes at 3 years of age.**

KEY WORDS: *cleft palate, speech outcome, timing of palatal surgery*

One of the factors that has been identified as influencing speech outcome for children with cleft palate is timing of primary palatal surgery. Supporters of early palatoplasty have stated that integrity of the speech mechanism is needed to enhance normal speech production and to minimize the development of compensatory articulations (CAs) (Peterson-Falzone et al., 2001). At the same time, opponents of this view have argued that early repair of the hard palate may have deleterious effects on midfacial growth (see Kuijpers-Jagtman and Long, 2000, for a review). Although this issue continues to be debated in the literature, both facial growth and speech considerations play an important role in clinical decisions regarding appropriate timing of cleft palate repair (e.g., Kemp-Fincham et al., 1990; Rohrich and Byrd, 1990; Rohrich et al., 1996, 2000; Nollet et al., 2005; Liao et al., 2006; Liao and Mars, 2006).

A majority of reports addressing timing of surgery have focused on the influence of age at time of palatal surgery on speech outcome (e.g., Holdsworth, 1954; Jolleys, 1954; McWilliams, 1960; Peet, 1961; Cleveland and Falk, 1970; Koberg and Koblin, 1973; Evans and Renfrew, 1974; Kaplan et al., 1974, 1980, 1982; Dorf and Curtin, 1982, 1990; Desai, 1983; Randall et al., 1983; Barimo et al.,1987; O'Gara and Logemann, 1988; Copeland, 1990; Rohrich and Byrd, 1990; Haapanen and Rantala, 1992; O'Gara et al., 1994; Denk and Magee, 1996; Rohrich et al., 1996, 2000; Marrinan et al., 1998; Ysunza et al., 1998; Kirschner et al., 2000; Sandberg et al., 2002; Rohrich and Gosman, 2004; Hardin-Jones and Jones, 2005).

Dr. Chapman is Associate Professor, Department of Communication Sciences and Disorders, University of Utah, Salt Lake City, Utah. Dr. Hardin-Jones is Professor, Division of Communication Disorders, University of Wyoming, Laramie, Wyoming. Dr. Goldstein is Chief of Plastic Surgery, Banner Children's Hospital; Medical Director, Craniofacial Center, Banner Children's Hospital, Mesa, Arizona; and is Professor, Department of Surgery, Case Western Reserve University, Cleveland, Ohio. Ms. Halter is Project Director, Regional Infant Hearing Program, Cleveland Hearing & Speech Center, Cleveland, Ohio. Dr. Havlik is Professor, Department of Surgery, Indiana University and Chief of Plastic Surgery, Riley Hospital, Indianapolis, Indiana. Ms. Schulte is Consultant, Indiana EHDI Program, Indiana State Department of Health, Indianapolis, Indiana.

This research was supported by a research grant from the National Institute on Deafness and Other Communication Disorders (RO1 DC03193).

Submitted December 2006; Accepted September 2007.

Address correspondence to: Kathy L. Chapman, Ph.D., Department of Communication Sciences and Disorders, 390 South 1530 East, University of Utah, Salt Lake City, UT 84103.

DOI: 10.1597/06-244.1

Despite major methodological differences and shortcomings of this literature (see Peterson-Falzone et al., 2001, for a detailed critique), a majority of studies suggested that earlier surgery was associated with better speech (Holdsworth, 1954; Jolleys, 1954; Peet, 1961; Cleveland and Falk, 1970; Evans and Renfrew, 1974; Dorf and Curtin, 1982, 1990; Randall et al., 1983; O'Gara and Logemann, 1988; O'Gara et al., 1994; Rohrich et al., 1996; Marrinan et al., 1998; Ysunza et al., 1998; Hardin-Jones and Jones, 2005). More specifically, early surgery was associated with better articulation (Evans and Renfrew, 1974; Kaplan et al., 1974; Dorf and Curtin, 1982, 1990; Rohrich et al., 1996; Ysunza et al., 1998), less CA usage (Dorf and Curtin, 1982, 1990; Ysunza et al., 1998), more normal resonance (Evans and Renfrew, 1974; Rohrich et al., 1996; Hardin-Jones and Jones, 2005), and/or less need of secondary surgery (Cleveland and Falk, 1970; Randall et al., 1983; Marrinan et al., 1998). Additionally, Kaplan et al. (1974) reported normal onset of babbling and more normal speech and language acquisition for children in their early surgery group compared with their late surgery group. Finally, others have reported a lower incidence of ear infections for children receiving early surgery (e.g., Kaplan et al., 1974; Randall et al., 1983).

One of many limitations of this literature is the lack of consensus on what constitutes early surgery (Peterson-Falzone et al., 2001). In fact, "early" surgery may have been performed anywhere from 1 day (Denk and Magee, 1996) to 2 years of age (Jolleys, 1954). Interestingly, although the independent variable in the studies described above was typically age at time of surgery, it was not chronological age *per se* that was at issue. Rather, it was often the case that the push for "early" surgery was related to the idea that surgery should be performed prior to the onset of speech/language (e.g., Wardill, 1933; Peet, 1961; Evans and Renfrew, 1974; Dorf and Curtin, 1982, 1990) and, thus, prior to 12 months of age (McWilliams, 1960; Peet, 1961). Although the proposal that surgery be performed prior to the onset of speech/language is not new (e.g., see Kaplan, 1981; Rohrich et al., 2000, for reviews), it was reinforced by the work of Dorf and Curtin (1982, 1990). Like many of the studies cited above, Dorf and Curtin were interested in comparing the speech outcome of children who received early and late surgery. Based on findings from 131 children, they determined that approximately 90% of children in the late surgery group (surgery after 12 months 15 days) exhibited CAs, compared with fewer than 5% of the children in the early surgery group (surgery prior to 12 months 15 days). More interestingly, however, they also found that the two children in the early surgery group who developed CAs experienced "early onset of phonemic development" (Dorf and Curtin, 1990, p. 348). At the same time, the children in the late surgery group who did not produce CAs had "slightly later than normal onset of phonemic development" (Dorf and Curtin, 1990, p. 348). The authors concluded that the child's age of phonemic development, rather than chronological age, should be considered when determining the optimum age for palatal surgery. More specifically, they argued that palatal surgery should be performed prior to the onset of meaningful speech. Since the original report was published in 1982, and despite methodological criticisms (see Kemp-Fincham et al., 1990; Peterson-Falzone, 1990; Chapman and Hardin, 1992; Dalston, 1992; Hardin-Jones and Jones, 2005), age at time of surgery in the United States has progressively decreased from 18 to 24 months of age in the 1970s to 11 months of age in the 1990s (Huebener and Marsh, 1997; Peterson-Falzone et al., 2001).

The idea that developmental criteria rather than chronological age should guide age of surgery decisions has strong theoretical support (see Kemp-Fincham et al., 1990), but empirical support is lacking. Only one study to date has attempted to examine the influence of lexical development at time of palatal surgery on speech outcome in general and on CA usage specifically. Chapman and Hardin (1992) examined the speech of children with cleft palate who received late surgery and were producing words when surgery was performed. The results were similar to Dorf and Curtin's (1990) in terms of the number of children producing CAs. That is, 90% of the children exhibited one or more instances of a CA. However, the median number of CAs was 6.5 with a range of 0 to 14. This is a very small frequency of occurrence considering that a total of 100 different words were transcribed for each child. Although Chapman and Hardin (1992) attempted to address the relationship between language status and speech outcome, their study had several limitations. First, although the children were producing words at the time of surgery, no specific information was available about vocabulary size nor did the study include a comparison group of children who were less advanced in terms of lexical development as a comparison.

The present study was designed to further this line of investigation by comparing the speech outcome of two groups of children with cleft palate who differed in lexical status at time of surgery. Because language skills and chronological age are correlated, it was typically the case that children who produced fewer words when surgery was performed were younger, and those who produced more words were older. We were interested in determining if children who were more lexically advanced and older at the time of primary palatal surgery had poorer articulation and resonance outcomes than those who were less lexically advanced and younger when primary palatal surgery was performed.

Method

Participants

Forty children with repaired cleft palate (with or without cleft lip) participated in the study. They ranged in age from

33 to 42 months (median age = 39 months) at the time of testing. The children were enrolled in a longitudinal study of speech and language development of children with cleft palate (with or without cleft lip) from 6 to 39 months of age. They were recruited from three craniofacial teams across the United States. All patients meeting the inclusion criteria during the enrollment period of the study were invited to participate. The study was approved by the Institutional Review Boards of the participating institutions, and informed consent was obtained for all participants. None of the children were diagnosed with syndromes, neurological impairment, or sensorineural hearing loss. Furthermore, all children were functioning within normal limits for cognitive skills (Bayley Scales of Infant Development-2 [BSID; Bayley, 1993]) and/or receptive language skills (Preschool Language Scale-3 [PLS-3; Zimmerman et al., 1992]).[1]

The children were recruited for the study between 3 and 6 months of age and were assigned to a group based on lexical development (i.e., size of vocabulary) when palatal surgery was performed (at 7 to 23 months of age). For three participants, surgery was attempted at approximately 12 months, but because the cleft was judged to be ‘‘wide,’’ only the soft palate was repaired and repair of the hard palate was delayed for a few months. These participants were assigned to groups based on vocabulary size at the time of hard palate repair. Group 1 (G1) included 20 children with less-advanced lexical development (i.e., fewer than five words reported by the mothers on the MacArthur Communicative Development Inventories [CDI; Fenson et al., 1993]) at the presurgery session. Group 2 (G2) included 20 children with more-advanced lexical development (five or more words reported by the mothers on the CDI) at the presurgery session (conducted within 2 weeks of surgery). Because language development and chronological age are related, children in G1 were typically younger than the children in G2 when palatal surgery was performed. So, although there was some overlap between the two groups in age of surgery, the G1 children had surgery at a mean age of 11 months (range = 7 to 14 months) compared with G2 children with a mean surgery age of 15 months (range = 12 to 23 months). Pairwise matching was employed; therefore, it was also the case that the child from G1 was younger than his/her match from G2 when surgery was performed (mean difference = 4 months; range = 1 to 10 months). The children also were matched for age at time of testing (±3 months), gender, race, and cleft type.[2] There was no significant difference between the groups in mother's level of education.

Because the participants were recruited from different sites across the United States, primary palatal surgery was performed by 10 different surgeons and type of surgery may have varied between and across surgeons. The children also had varying intervention histories, with 43% (17/40) of the children being enrolled in intervention at time of testing. All but one of the study participants had pressure equalization (PE) tubes inserted either prior to or at the time of palatal surgery (typically at the time of lip repair). The median age for initial PE tube insertion was 6 months (range = 2 to 11 months) for the children in G1 (< five words/earlier surgery) and 5 months (range = 2 to 20 months) for the children in G2 (≥ five words/later surgery). Thirty-seven (out of 40) participants had at least one reported episode of otitis media over the period of the study (based on parent report and an examination of medical records). Finally, five children had small palatal fistulas (three anterior and two posterior) that did not appear to be symptomatic for speech. An additional child had a fistula that was symptomatic for speech; however, that child received surgical repair of the fistula prior to the time of testing for this study. (Please see Table 1 for additional information concerning participant characteristics.)

Data Collection

All data collection sessions were conducted in the children's homes. A majority of the children were 39 months at the time of testing. However, in some cases, because this was a longitudinal investigation, children may have missed a session (or the session was delayed) due to illness, a family move, or other unavoidable circumstances. In those cases, the 33-month session was analyzed for that child and his/her matched pair. At the 33-month session, in addition to a spontaneous speech sample (see below), the Goldman-Fristoe Test of Articulation (Goldman and Fristoe, 1986) was administered. At the 39-month session, the PLS-3 and the BSID were added to the experimental testing. At the end of all sessions, information was obtained from caregivers concerning the children's middle ear status and pertinent management history. Finally, tympanometry was performed using a Micro Audiometrics Earscan (Model 2112; Murphy, NC).

[1] Two of the children (a matched pair) exhibited cognitive skills within normal limits at an earlier age but were below normal limits on the Mental scale of the BSID for at least one of the later testing sessions. However, they scored within normal limits on the receptive and expressive subtests of the PLS at all testing points. Because their performance on the BSID at the later experimental session may have been influenced by social and cultural factors (e.g., low socioeconomic status, low maternal education level) and their performance on the language test was age-appropriate, they were included in this study.

[2] Two of the matched pairs included a girl with a unilateral cleft lip and palate being matched to a girl with a bilateral cleft lip and palate. This was done only after a comparison was made including all children in the larger longitudinal study based on cleft type (unilateral cleft palate versus bilateral cleft palate). Results of multiple t tests suggested no differences between the two cleft groups for the independent variables of interest in this study.

TABLE 1 Participant Characteristics for Children in Group 1 (< Five Words/Earlier Surgery) and Group 2 (≥ Five Words/ Later Surgery)*

	Group 1	*Group 2*	*Total*
Number of children			
Boys	10	10	20
Girls	10	10	20
Type of cleft			
BCLP	3	5	8
UCLP	13	11	24
HSP	2	2	4
SPO	2	2	4
No. of words at time of surgery			
Median	1	8.5	
Range	0 to 4	5 to 90	
Age at time of surgery			
Median (mo)	11	14	
Range (mo)	7 to 14	12 to 23	
Age at first PE tube insertion			
Median (mo)	6	5	
Range (mo)	2 to 11	2 to 20	
No. of ear infections at time of testing			
Median	5	7	
Range	0 to 16	0 to 16	
No. of children enrolled in intervention at time of testing	5	12	17

* BCLP = bilateral cleft lip and palate; UCLP = unilateral cleft lip and palate; HSP = cleft of the hard and soft palates; SPO = cleft of the soft palate only; PE = pressure equalization.

The spontaneous speech sampling portion of the session lasted approximately 45 minutes or until the child produced at least 400 words. Each child interacted with his/her caregiver using a variety of age-appropriate toys. The toys were chosen specifically to sample all sounds, across all word positions. Caregivers were instructed to interact with their children as they normally would. All sampling sessions were audio- and video-recorded with a Countryman wireless microphone (MEMWF5ETS; McKee Rocks, PA), Telex receiver (RMR-70) and transmitter (WT60; Garland, TX), Marantz portable cassette recorder (PMD 430; Chatsworth, CA), and Panasonic video camera (Model AG188; Secaucus, NJ).

Data Analysis

Speech outcome for the two groups of children was examined in two ways. First, measures of speech outcome (described in detail below) were obtained from the spontaneous speech samples collected at the sampling sessions. Second, a 30-second sample of each child's spontaneous speech (also obtained from the sampling sessions) was rated by a group of listeners for articulation proficiency and hypernasality, separately, using direct magnitude estimation (DME).

Spontaneous Speech Sample Data

The audio-recorded spontaneous speech samples were transcribed independently by K. Chapman or M. Hardin-Jones, speech-language pathologists with more than 20 years of experience transcribing the speech of babies and young children with cleft palate. The samples were transcribed in a quiet room using the Marantz recorder (see above) and external speakers (Bose Lifestyle Powered Speakers; Framingham, MA). Samples were collected over a period of several years, and transcriptions were performed over that same period of time. However, the investigators met approximately every 6 months to perform reliability checks to guard against transcriber drift. When needed, the video recordings were reviewed to obtain glosses for unintelligible words. The International Phonetic Alphabet (International Phonetic Association, 1999) and CA symbols (Trost, 1981) were employed for transcription of the samples. Two hundred words (including at least 100 different words) were transcribed for each child. The transcribed samples were entered on computer and were analyzed using Logical International Phonetics Programs (Oller and Delgado, 2000).

A number of measures were computed from the spontaneous speech data, including size of consonant inventory, size of true consonant inventory (i.e., all consonants except glides, glottals, laryngeals, or pharyngeals), number of stable consonants, total consonants correct, percentage correct for manner of articulation categories (i.e., percent correct for stops, nasals, fricatives, affricates, liquids, and glides), and percentage occurrence for CAs. The two consonant inventory measures were made by counting all consonants (size of consonant inventory) and all true consonants (size of true consonant inventory) that occurred at least two times in the sample. Stable consonants included all sounds that were correct at least 70% of the time. Total consonants correct was calculated by dividing the number of correct consonants by the total number of consonants produced. Percentage occurrence for each of the manner of articulation categories was calculated by dividing the total number of correct productions by the total number of sounds produced for each manner category (e.g., number of correct stops / total number of stops). Speech status was rated as (1) age appropriate, (2) speech delayed, or (3) speech delayed with CAs. Normative data of Smit et al. (1990) were used to determine if a child's speech was age appropriate or delayed. Children who produced more than two age-appropriate sounds with less than 75% accuracy were considered to be speech delayed.[3] Finally, the two transcribers rated each child's spontaneous sample for degree of hypernasality. Samples were rated as normal, mildly hypernasal, moderately hypernasal, severely hypernasal, or hyponasal. Significant hypernasality was considered to be present for children who received a rating of

[3] Examination of the speech results for our noncleft participants indicated that very few (<10%) produced /t/ with 75% accuracy; although, Smit et al. (1990) found that /t/ was acquired by 75% of the children in her study by age 3. Therefore, in this study, /t/ was not considered to be a sound that should have been acquired by age 3.

moderate or severe hypernasality. VP status was considered to be inadequate if the child had undergone secondary surgery or if hypernasality was rated as moderate or severe.

Descriptive statistics were computed for the dependent variables of interest (see Table 2). Paired sample *t* tests were used to determine if significant differences existed between the two groups of children with cleft palate on any of these variables. Chi-square analyses were performed to examine the relationship between group assignment and a number of outcome variables including speech status, VP status, and intervention status (i.e., whether or not the child was enrolled in intervention at the time of testing).

DME Ratings of Articulation Proficiency and Resonance Outcome

To determine if differences existed between the two groups of children with cleft palate in articulation proficiency (i.e., age-appropriate articulation and phonology) and/or hypernasality, ratings of the children's connected speech at 33 or 39 months of age were performed by 10 experienced speech-language pathologists who were blinded to the children's group assignment. The ratings were completed using DME with a modulus (description of the procedure can be found below). Recent investigations have suggested that many judgments of speech and/or resonance should not be performed using equal-appearing interval (EAI) scaling as it may be difficult for listeners to perform these ratings in an unbiased manner (e.g., Schiavetti et al., 1981, 1983; Zraick and Liss, 2000; Eadie and Doyle, 2002; Whitehill et al., 2002). This is true for dimensions that are prothetic such as loudness (i.e., "continua concerned with intensity or amount" [Stevens, 1975, p. 13]) because results of experiments reported by Stevens (1975) suggested that "listeners subdivide the lower end of the scale into smaller intervals than the upper portion of the continuum" (Whitehill et al., 2002, p. 81). Results of a recent study by Whitehill et al. (2002) comparing EAI and DME scaling of hypernasality concluded that because hypernasality is also prothetic, EAI scaling "may not be appropriate," but DME was "a valid and reliable measure" for judging hypernasality in individuals with cleft palate (Whitehill et al., 2002, p. 85).

The speech samples to be rated consisted of 30 seconds of spontaneous speech from each child's 33- or 39-month spontaneous speech sample (samples from two participants can be found in Appendix A). Samples were chosen that contained the longest stream of uninterrupted child speech. Attempts were made during the sampling sessions to elicit spontaneous personal narratives from the children, however, this was not always successful. Samples were also included from typically developing 3-year-olds without cleft palate who were participating in the longitudinal study described above. Due to the young age of the children, the samples varied in content as the children were unable to read or to repeat utterances as long as 30 seconds.

A modulus or standard sample was chosen for each rating task (i.e., one modulus for articulation proficiency and one modulus for hypernasality) from samples of connected speech produced by children of the same age who also were participating in the longitudinal study but were not included in the experimental task described here. These samples were chosen and prepared in the same manner as the experimental samples. The samples were independently reviewed by the two investigators (K. Chapman and M. Hardin-Jones) and both agreed on a modulus sample for articulation proficiency and hypernasality, respectively. Each was considered to be the best example of the "midpoint" for articulation proficiency and hypernasality of study participants. Although DME ratings can be performed without a modulus, many judges are less comfortable with this procedure, and the ratings must be normalized (different raters apply a different range of scores) prior to data analysis (Weismer and Laures, 2002).

The samples were digitized at a sampling rate of 50 kHz using Computerized Speech Lab (CSL 4100; KayPEN-

TABLE 2 **Means (M), Standard Deviations (SD), *t* Values (*t*), Significance Levels (*p*), and Effect Sizes (*d*) for the Speech Outcome Measures for Children in Group 1 (< Five Words/Earlier Surgery) and Group 2 (≥ Five Words/Later Surgery)**

	Group 1		*Group 2*				
Measures	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*
Size of consonant inventory	19.15	2.16	17.50	3.29	2.07	.05*	.59
Size of true consonant inventory	15.20	2.31	13.55	3.33	2.08	.05*	.57
No. of stable consonants	10.40	4.14	9.25	3.39	1.24	.23	.30
Total consonants correct	61.43	14.01	55.53	15.50	1.52	.15	.40
% stops correct	58.68	19.08	53.59	22.20	0.94	.36	.25
% nasals correct	82.95	10.40	76.97	10.71	2.58	.02*	.57
% fricatives correct	54.96	16.03	47.21	17.67	1.48	.16	.45
% affricates correct	39.34	29.46	25.92	29.02	1.76	.09	.37
% liquids correct	32.90	25.08	20.26	13.91	2.06	.05*	.62
% glides correct	89.59	11.99	91.80	6.29	−0.93	.36	.23
% CA productions†	3.56	5.97	4.18	7.87	−1.33	.20	.38

† CA= compensatory articulation.

* Significant at $p = .05$; Cohen's effect size: small = .20, medium = .50, large =.80 (Cohen, 1988).

TAX, Lincoln Park, NJ) and a Dell Dimensions T600 series computer (Round Rock, TX). Digitized samples were saved as WAV files and Adobe Audition software was used to edit the samples. Background noise and distracting nonspeech artifacts (e.g., toy noise) were attenuated using the software's "noise reduction" and "envelope" options. Next, in order to achieve a consistent amplitude on the listening CDs, the loudness of individual audio-recorded samples was "normalized" to 100% using the Adobe Audition software. Introductions were added to identify the order of sample presentation on the listening CD, and a 5-second interval of silence was appended to the end of each sample. Edited files were burned to a CD in random order. Thirteen randomly chosen samples (approximately 30% of the samples) were repeated for reliability purposes, and four samples representing the speech of children without cleft palate were included in random order. Thus, each of the listening tasks (articulation proficiency and hypernasality) consisted of 57 experimental samples (40 children with cleft palate, four noncleft children, and 13 reliability samples) and 19 modulus samples (presented after every third experimental sample). Two sets of training samples (one for articulation proficiency and one for hypernasality) were compiled in the same fashion. The majority of these samples were obtained from children with cleft palate who also were participating in the larger longitudinal study and represented a range of values, comparable to the experimental samples, in terms of articulation proficiency and hypernasality. Each training task consisted of nine samples, and the modulus was presented after every third sample.

The experimental samples (57 total) were independently rated for (1) articulation proficiency and (2) hypernasality by 10 certified speech-language pathologists (the listeners) with at least 5 years of experience (median number of years = 16.5; range = 6 to 40 years) listening to young children with severe speech problems. One half of the listeners rated articulation proficiency first and the other half rated hypernasality first. The listeners were taught to use DME scaling with a modulus to rate the samples (e.g., Eadie and Doyle, 2002; Whitehill et al., 2002). For judgments of articulation proficiency, the listeners were instructed to consider the child's articulation/phonology and to ignore hypernasality when assigning an overall rating. When rating hypernasality, the listeners were instructed to consider the degree of nasal resonance but to ignore articulation/phonology. At the start of each rating task, the listeners were presented with the modulus sample for the dimension of interest and were told that it represented a value of 100 on the scale. Listeners were told to rate all additional samples with reference to the modulus sample. For example, when rating hypernasality, if a sample was thought to be twice as hypernasal as the modulus sample presented, the listeners assigned a rating of 200. Similarly, if the sample was believed to be about half as hypernasal as the modulus, the listeners assigned a rating of 50. The same procedure was employed for rating articulation proficiency. However, a sample that was judged to be "twice as good" as the modulus received a rating of 200, and a sample judged to be about "half as good" received a rating of 50 (Eadie and Doyle, 2002). The experimental samples were preceded by the nine training samples for practice with the DME rating scale. The samples were presented using a laptop computer and Sony earphones (MDR7506; New York, NY) at a comfortable listening level in a sound-treated room.

Finally, separate independent sample t tests were performed to determine if there were differences in the mean DME ratings for articulation proficiency and hypernasality for the two groups of children with cleft palate. Based on results of the intrajudge reliability measure (see section on reliability below), the resonance ratings for one listener and the articulation proficiency ratings for another listener were omitted from the statistical analyses. Their scores fell below .70, which was considered to be the cutoff for acceptable levels of intrajudge reliability. Prior to statistical analysis, the DME ratings were transformed using a base 10 logarithm (lg10) (SPSS Version 14, SPSS Inc., Chicago, IL), and t tests were calculated using geometric means. Use of the geometric mean rather than the arithmetic mean as the measure of central tendency is suggested for ratings obtained using DME procedures. According to Marks (1974) and Stevens (1975), the geometric mean is employed because it is less affected by high scores and provides "an unbiased estimate of the expected value of the logarithms of the magnitude estimates" (Marks, 1974, p. 44).

Reliability

Spontaneous Speech Sample Data

Six randomly chosen samples were transcribed independently by a second transcriber (K. Chapman or M. Hardin-Jones) to calculate interjudge reliability. Only consonants that were transcribed identically for place, manner, and voicing were considered to be agreements. The percentage agreement scores were calculated by dividing the number of agreements by the number of consonants produced in the sample (i.e., agreements and disagreements). Interjudge reliability ranged from 78% to 91% (mean = 85%).

DME Ratings of Articulation Proficiency and Resonance Outcome

Reliability also was obtained for the listeners' ratings of articulation proficiency and hypernasality. Intrajudge reliability was calculated for each listener on the 13 reliability tokens using Pearson product moment correlations for both articulation proficiency and hypernasality, separately. The intrajudge reliability coefficients were r = .88 for speech proficiency and r = .87 for hypernasality;

they ranged from r = .85 to .96 for speech proficiency and from r = .71 to .96 for hypernasality. The Cronbach alpha (interclass correlation type 3,k) (SPSS Version 14.0) was used to calculate interjudge reliability. See Shrout and Fleiss (1979) for a complete description of interclass correlations. Group reliability coefficients were .96 for articulation proficiency and .95 for hypernasality.

Results

Spontaneous Speech Sample Data

The first statistical tests compared the two groups of children with cleft palate on several measures of speech production based on spontaneous speech sampling. Results of multiple paired sample t tests revealed statistically significant differences between the two groups of children for size of consonant inventory ($t = 2.07$; $p = .05$), size of true consonant inventory ($t = 2.08$; $p = .05$), percentage of correct nasals ($t = 2.58$; $p = .02$), and percentage of correct liquids ($t = 2.06$; $p = .05$). In all cases, the children in G1 (< five words/earlier surgery) performed better than the children in G2 (≥ five words/later surgery). All other comparisons were nonsignificant (see Table 2). Effect sizes (Cohen's d [Cohen, 1988]) were calculated following the procedures described by Dunst et al. (2004) for nonindependent sample designs. Values ranged from .57 to .62 for the significant comparisons, indicating medium effect sizes (see Table 2).

The second set of statistical tests examined the relationship between group assignment and a number of outcome variables. Chi-square analyses indicated a relationship between group assignment and the number of children enrolled in intervention at the time of testing ($\chi^2 = 5.01$; $p = .03$). In this case, only 25% (5/20) of the children in G1 (< five words/earlier surgery) were enrolled in intervention compared with 60% (12/20) of the children in G2 (≥ five words/later surgery). However, all other chi-square tests were nonsignificant. For example, no relationship was found between group assignment and the number of children requiring secondary surgery or with significant hypernasality ($\chi^2 = .10$; $p = .75$). Forty-five percent (9/20) of the children in G1 (< five words/earlier surgery) received secondary surgery or exhibited significant hypernasality compared with 40% (8/20) in G2 (≥ five words/later surgery). Similarly, no significant relationship was noted between lexical status/age at time of surgery and number of children with normal speech development ($\chi^2 = .96$; $p = .33$). In this case, 45% (9/20) of the children in G1 (< five words/earlier surgery) had normal speech compared with 30% (6/20) in G2 (≥ five words/later surgery).

DME Ratings of Articulation Proficiency and Resonance Outcome

The final statistical analyses compared the DME ratings for the two groups of children with cleft palate for articulation proficiency and hypernasality using separate independent sample t tests. Significant differences were found for articulation proficiency ($t = 3.74$; $p = .01$) and hypernasality ($t = -2.14$; $p = .03$). The children in G1 (< five words/earlier surgery) received a higher mean rating for articulation proficiency (M = 120.06) compared with children in G2 (≥ five words/later surgery) (M = 86.86), suggesting better articulation skills for participants who were less lexically advanced and younger at the time of surgery. The children in G2 (≥ five words/later surgery) received a higher mean rating for hypernasality (M = 97.08) than children in the G1 (< five words/earlier surgery) (M = 81.64). However, in this case a higher score was associated with a greater degree of nasal resonance and a poorer outcome. These findings suggest better speech/articulation and resonance for the children with cleft palate who were younger and less lexically advanced at the time of palatal surgery.

Discussion

The purpose of the present study was to compare the speech outcome of two groups of children with cleft palate—those who were younger and less lexically advanced and those who were older and more lexically advanced at the time of primary palatal surgery. More similarities than differences were noted from the analyses of the spontaneous speech sample data. For example, the two groups of children were similar in their production of stable consonants, total consonants correct, stops, fricatives, affricates, glides, and compensatory productions. Additionally, no relationship was noted between lexical status/age at time of surgery and either VP status or speech status. At the same time, differences were observed between the groups for several outcome measures. Children in G1 (< five words/earlier surgery) had larger consonant inventories and produced more correct nasals and liquids. It was also the case that children in G1 were less likely to be enrolled in intervention at 3 years of age. Interestingly, the listeners assigned higher articulation proficiency scores and lower hypernasality scores to this group of children. This latter finding suggested that, overall, children who were less lexically advanced and younger at the time of surgery (i.e., G1) were judged to have a better speech outcome than their peers who were more lexically advanced and older when palatal surgery was performed (i.e., G2).

At first glance, the results obtained from the DME ratings appear inconsistent with the results obtained from the children's spontaneous speech sample data. However, an examination of the mean scores for the consonant production variables (Table 2) showed that the children in G1 (< five words/earlier surgery) typically showed higher scores than the children in G2 (≥ five words/later surgery), suggesting better performance overall. Moreover, more of the children in G1 were rated as having "normal" speech sound development (45% of the children in G1 compared with 30% in G2). Additionally, examination of the resonance ratings assigned by the authors (see "Data Analysis" section for a description) showed that although the number of children with and without significant hypernasality was very similar for the two groups, almost twice as many of the children in G1 (< five words/earlier surgery) were rated as having normal resonance (45% versus 25% of the children in G1 and G2, respectively). The additive effect of all of these differences may have been revealed in the DME ratings, suggesting better speech and resonance outcomes for children who were younger and less lexically advanced at the time of surgery.

Recall that in the present study, the DME procedure required the listeners to engage in separate rating tasks for articulation and resonance (and they were instructed to focus on one dimension while ignoring the other). It seems likely, however, that these two aspects of speech influenced each other to some degree because many of the children who exhibited hypernasality secondary to velopharyngeal inadequacy (VPI) tended to have concomitant articulation problems. To examine this interaction, we performed a correlation between the scores obtained for articulation proficiency and hypernasality for the two groups of children combined. A significant ($p < .01$) but moderate correlation of $-.41$ was noted between the two variables (Pearson product moment correlation), suggesting that higher (better) ratings of articulation proficiency were associated with lower ratings of hypernasality (and *vice versa*). Additionally, we had employed objective criteria to evaluate each child's speech skills based on accuracy of production for age-appropriate sounds (see "Data Analysis" section). Therefore, we were able to compare the mean DME articulation proficiency rating with the objective rating of speech status (i.e., age appropriate or speech delayed) each child was assigned based on his/her sound production accuracy from the spontaneous speech sample data. There were only four cases where the mean DME articulation proficiency scores assigned by the listeners appeared to be lower than would be expected based on the ratings of speech status. For two of the participants (one in G1 and one in G2), although they both exhibited good speech production skills, they also exhibited moderate hypernasality, which may have influenced the scores assigned by the listeners. For the other two children, hypernasality did not appear related to their lower ratings on the DME task. Unfortunately, the reciprocal relationship between poor speech and judgments of hypernasality is less easy to evaluate because we did not have an objective measure of hypernasality for comparison. However, regardless of the effect that hypernasality may have had on judgments of articulation and *vice versa*, the children who were more lexically advanced and older at time of surgery were judged to have poorer DME ratings on both dimensions.

The results of this study provide support for the use of DME scaling by experienced speech-language pathologists to evaluate both articulation proficiency and resonance of children with cleft palate. Interestingly, the listeners appeared to be responding to differences in speech that were not quantified as easily when comparing the groups for accuracy of production across a number of consonant production features. More research should be undertaken that employs DME scaling for measuring speech outcome of children with cleft palate. For example, it would be interesting to examine how the different measures of articulation employed in this study (e.g., total consonants correct, number of stable consonants, percentage of correct stops) were more or less related to the ratings of articulation proficiency and how they interacted with hypernasality to influence overall intelligibility. Such information is important for understanding how the various components of speech/resonance contribute to intelligibility for all speakers, including those with cleft palate, and also has implications for developing priorities for intervention (Whitehill and Chun, 2002). Finally, "standard" samples for speech and hypernasality could be identified for speakers of different ages, thus facilitating comparisons across data sets, investigators, and treatment centers.

There are several possible explanations for the finding of a better speech outcome for the children who were younger and less lexically advanced at the time of surgery. First, because these children were younger when surgery was performed, they had more time to "practice" with a repaired mechanism. The role that lexical status may have played in the outcome is less straightforward. There was no difference between the two groups in CA usage, suggesting that more-advanced lexical development at the time of palatal surgery did not result in greater CA usage. Based on what is known about early phonological development in typically developing children, these findings are not too surprising. Many believe that during the earliest period of word-learning, children are targeting "whole-word patterns" not individual sounds. It is not until children have a vocabulary of approximately 50 words that the phonological system becomes systematic (i.e., rule-governed) (see Kemp-Fincham et al., 1990; Vihman, 2004, for reviews). Only one of the children in G2 (≥ five words/later surgery) had reached that point at the time of palatal surgery. The other 19 had expressive vocabularies ranging from 5 to 38 words (the one child was combining words and had an expressive vocabulary of 90 words). Moreover, a careful inspection of our data suggests that CA usage was related to VP status rather than to group assignment. That is, seven of the nine children (five children from G1 and four from G2) who produced CAs eventually received secondary surgery or exhibited moderate or severe hypernasality and other speech behaviors characteristic of VPI. On that basis, we might hypothesize that CAs are likely to be seen in the speech of children who produce more than 50 words (18 to 24 months of age) at the time of palatal surgery or who are younger and less lexically advanced at the time of surgery but who continue to acquire speech/language in the presence of VPI (Hardin-Jones and Jones, 2005).

Another explanation that must be considered is that the poorer results for children who were more lexically advanced and older at the time of palatal surgery (G2) may have been related to the fact that this group contained more children with severe clefts. This could only have been the case, however, if surgery was routinely postponed for children based on extent of the cleft and if a disproportionate number of these children were assigned to G2 (≥ five words/later surgery). The surgeons on our teams typically performed surgery at approximately the same age for all children with nonsyndromic cleft palate (with delays related to illness or scheduling issues). There were even a few cases where surgery was attempted at approximately 12 months, but because the cleft was judged to be "wide," only the soft palate was repaired and closure of the hard palate was delayed for a few months. Although these children were all assigned to G2 (≥ five words/later surgery), they did not have a significantly worse outcome. Also, recall that the children in this study were assigned to the groups based on number of words produced rather than age at time of surgery. So, whereas the children in G2 were usually older and those in G1 were typically younger, this was not always the case. There were eight children who produced more than five words but had surgery before 12 months (before 12 months 15 days) or had fewer than five words but had surgery after 12 months.

Unfortunately, other than cleft type, we did not have documentation of extent of clefting for all child participants. We performed some preliminary analyses comparing the speech characteristics of children with bilateral versus unilateral clefts in our entire data set, as well as for final study participants. Interestingly, there were no significant differences between these two groups for any of the speech outcome variables based on cleft type. In fact, a higher percentage of children with bilateral cleft lip and palate had speech development within normal limits (50% versus 33% for children with unilateral cleft lip and palate) and adequate velopharyngeal function (63% versus 54% for children with unilateral cleft lip and palate). Additionally, there were no significant differences in the DME ratings for children with unilateral versus bilateral cleft palate. Finally, information on cleft width was available for 15 children in our data set (child participants in Indiana). Thus, a correlation was performed to determine if a relationship existed between cleft width and speech or VP status, between cleft width and cleft type, and between cleft width and group assignment. None of the correlations were significant, all were low, and two were negative. Additional research is needed to determine if there are variables other than cleft type or cleft width that might need to be considered when examining speech outcome for children with cleft palate (see Mølsted, 1999, for a discussion of variability in clefting and what some of these measures might be).

Finally, one might question whether some of the differences between the groups could be related to fewer middle ear problems and less conductive hearing loss in children who received earlier surgery (i.e., G1). Although we did not have data that would allow us to compare the hearing histories of the two groups of children, we did have information concerning number of episodes of otitis media over the study period. No significant differences were noted between the groups for number of reported episodes of otitis media, suggesting that the differences reported here were probably not the result of differences in otologic/audiologic status between the two groups.

A challenging aspect of the current study design was determining the criteria for assigning the children with cleft palate to the two groups. Unfortunately, we were not able to replicate what was done by Dorf and Curtin (1990), because they did not describe what they meant by "early onset of phonemic development" or "articulation age." In the current study, participants were divided into groups based on size of lexicon (G1 = fewer than five words; G2 = five words or more) obtained from parent report of number of words produced on the CDI and examination of a spontaneous speech sample, both obtained within 2 weeks of primary palatal surgery. Either was taken as proof of word usage. Of course, there is a possibility that the findings would have been different if another cutoff had been used for assigning children to the groups. For example, we might have used no words produced versus one word produced for separating the two groups. Or, employing the guidelines proposed by Vihman (1986), a child would have had to produce four or more different words in a session to be considered linguistic, regardless of information about word use that was supplied by the parents. Vihman's criteria seemed too stringent as it was not always the case that children who were described by parents as having five words always showed production of these words during the sampling session. Previous work suggested that looking only at word usage during sampling would have underestimated vocabulary size in our children (Chapman et al., 2003). Additionally, the CDI has proven to be a valid and reliable measure of expressive vocabulary for typically developing children (see Dale, 1996, for a review) as well as for children with cleft palate (Scherer and D'Antonio, 1995). Finally, an examination of the child pairs showed a clear difference in the number of words produced for those children in G1 (< five words/earlier surgery) compared with those in G2 (≥ five words/later surgery). It was not the case that the children were separated by a few words. For example, children who were in G2 but were on the lower end in terms of number of words produced (e.g., those children who produced five words) were paired with children who produced no words. Or, children in G1 who were at the higher end (i.e., produced 3 to 4 words) were matched with children from G2 who produced at least 10 words. Therefore, the groups were more different from each other in terms of lexical development than the cutoff would have implied.

The findings of this study have implications for current views on timing of palatal surgery. Craniofacial team members might want to reconsider the view that palatal surgery performed by 12 months of age and prior to the onset of words is early enough. Although the work of Dorf and Curtin (1982, 1990) has likely played an important role in decreasing the age of palatal surgery in the United States, it also may have been used as a rationale for not performing surgery earlier than 12 months. The results of this study suggest that early surgery is better from a speech and resonance perspective; however, it is not clear how early surgery needs to be to ensure the best possible outcome. These findings suggest that 11 months is better than 15 months, but is earlier than 11 months even better? Based on studies of surgery performed earlier than 11 months (see introductory paragraphs) and from information gleaned from the normal speech and language development literature, we hypothesize that performing surgery by 6 months may result in an even better speech outcome. Although primary palatal surgery prior to 11 months of age will provide most babies with an adequate mechanism before the onset of words, speech development does not begin with production of the first word. Rather, early prelinguistic sound productions or babbling provides the foundation for early word productions (e.g., Vihman, 1992; Stoel-Gammon, 1992, 1998a, 1998b). Investigations of the babbling of babies with cleft palate suggest that they also produce a variety of sounds/syllables in their early babbling. However, their early babbling has been found to differ from that of noncleft babies in the types of sounds that are produced and frequency of canonical syllables (well-formed syllables that resemble adult speech [Oller, 2000]). Babies with cleft palate also may enter the canonical babbling stage later than noncleft babies (Chapman et al., 2001). Therefore, it seems important that surgery be performed as early as possible but at least prior to the onset of canonical babbling. Although data are not available on the actual age of onset of canonical babbling for babies with cleft palate, we know that approximately 50% of babies exhibit canonical babbling by 9 months of age (Chapman et al., 2001). Thus, surgery by 6 months would ensure that palatal repair was performed prior to that time.

Finally, the work by Dorf and Curtin (1982, 1990) may have been used by some as a rationale for suggesting that parents not stimulate their baby's language until after primary palatal surgery in an attempt to prevent the development of CAs. Given the typical age for primary palatal surgery in the US, it is unlikely that many babies will be beyond the 50 word stage when surgery is performed. Even so, babies with cleft palate should receive the same early language stimulation as noncleft babies regardless of whether or not the palate has been repaired if normal speech and language development is the goal.

Although the present study attempted to overcome some of the shortcomings of previous research examining the impact of timing of surgery on speech outcome, it also has some limitations that must be acknowledged. First, by the time the children with cleft palate were matched on important variables that also might influence outcome, there were only 20 children per group, although 62 children with cleft palate were enrolled in the longitudinal study. There is a possibility that with a larger sample, more differences between the groups would have emerged in favor of those children who were younger and less lexically advanced at the time of surgery.

Second, although this was a prospective, multicenter, longitudinal study that employed pairwise matching and blinded raters, it was a descriptive study. As a result, there is always the possibility of selection bias and confounding variables influencing the findings (Stirrat et al., 1992). Because of these and other shortcomings of descriptive studies, some have suggested that randomized controlled clinical trials (RCTs) are the appropriate next step for examining questions related to surgical outcome in the area of cleft palate (Semb and Shaw, 1998; Mølsted, 1999). At the same time, others believe that RCTs are not employed as easily to evaluate surgical techniques in general (Fung and Loré, 2002) and cleft palate surgery specifically (Berkowitz, 1995). The particular question we were interested in here could not have been addressed with a RCT because group assignment was based on language level at time of surgery. As a result, random assignment of the participants was not possible. However, because this was a descriptive study, the limitations associated with such studies must be considered when interpreting the results.

Finally, this study only examined one outcome, speech, and did not consider others such as facial growth and appearance, occlusion, or burden of care that are also important when planning treatment for children with cleft palate.

In conclusion, the findings of this study support previous research showing that a better speech outcome was associated with earlier surgery. At the same time, children who were older and had larger vocabularies prior to surgical repair of the palate did not produce CAs at a greater rate than children who were younger and had less-advanced vocabulary development (although the former group exhibited poorer speech in general). Future studies should compare children undergoing surgery by 6 months or prior to the onset of babbling with children receiving surgery at approximately 12 months. Other outcomes such as facial growth, appearance, and occlusion also should be evaluated along with speech in the same studies to determine the impact of the timing decisions on all aspects of treatment outcome. Because of improvements in surgical and behavioral interventions and the large variation across centers in timing of palatal surgery, questions related to timing of surgical intervention continue to be an important area of study.

Acknowledgments. This work was supported by a research grant from the National Institute on Deafness and Other Communication Disorders (R01

DC03193). We wish to thank the parents and children who participated in this study. We also thank Katie Deninger, Rebecca Fenander, Melissa Frederickson, Courtney Jones, Amy Krantz, Marilyn Watson, and Jane Wright for their assistance with various aspects of this study.

References

Barimo JP, Habal MB, Scheuerle J, Ritterman SI. Postnatal palatoplasty, implications for normal speech articulation—a preliminary report. *Scand J Plast Reconstr Surg Hand Surg*. 1987;21:139–143.

Bayley N. *Bayley Scales of Infant Development*. 2nd ed. San Antonio: The Psychological Corporation; 1993.

Berkowitz S. Ethical issues in the case of surgical repair of cleft palate. *Cleft Palate Craniofac J*. 1995;32:271–281.

Chapman KL, Hardin MA. Phonetic and phonological skills of two-year-olds with cleft palate. *Cleft Palate Craniofac J*. 1992;29:435–443.

Chapman KL, Hardin-Jones M, Halter KA. The relationship between early speech and later speech and language performance for children with cleft palate. *Clin Linguist Phon*. 2003;17:173–197.

Chapman KL, Hardin-Jones M, Schulte J, Halter K. Vocal development of 9-month-old babies with cleft palate. *J Speech Hear Res*. 2001;44:1268–1283.

Cleveland KM, Falk ML. Several factors which may precipitate the use of pharyngeal flap. *Cleft Palate J*. 1970;7:105–111.

Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

Copeland M. The effects of very early palatal repair on speech. *Br J Plast Surg*. 1990;43:676–682.

Dale PS. Parent report assessment. In: Cole KN, Dale PS, Thal DJ, eds. *Assessment of Communication and Language*. Vol. 6. Baltimore, MD: Paul H. Brookes Publishing; 1996:161–182.

Dalston RM. Timing of cleft palate repair: a speech pathologist's viewpoint. *Probl Plast Reconstr Surg*. 1992;2:30–38.

Denk MJ, Magee WP. Cleft palate closure in the neonate: preliminary report. *Cleft Palate Craniofac J*. 1996;33:57–61.

Desai SN. Early cleft palate repair completed before the age of 16 weeks: observations on a personal series of 100 children. *Br J Plast Surg*. 1983;36:300–304.

Dorf DS, Curtin JW. Early cleft palate repair and speech outcome. *Plast Reconstr Surg*. 1982;70:74–79.

Dorf DS, Curtin JW. Early cleft palate repair and speech outcome: a ten-year experience. In: Bardach J, Morris HL, eds. *Multidisciplinary Management of Cleft Lip and Palate*. Philadelphia: WB Saunders; 1990:341–348.

Dunst CJ, Hamby DW, Trivette CM. Guidelines for calculating effect sizes for practice-based research syntheses. *Centerscope*. 2004;2:1–9.

Eadie TL, Doyle PC. Direct magnitude estimation and interval scaling of pleasantness and severity in dysphonic and normal speakers. *J Acoust Soc Am*. 2002;112:3014–3021.

Evans D, Renfrew C. The timing of primary cleft palate repair. *Scand J Plast Reconstr Surg*. 1974;8:153–155.

Fenson L, Dale PS, Reznick JS, Thal D, Bates E, Hartung JP, Pethick S, Reilly JS. *MacArthur Communicative Development Inventories*. San Diego, CA: Singular Publishing Group; 1993.

Fung EK, Loré JM. Randomized controlled trials for evaluating surgical questions. *Arch Otolaryngol Head Neck Surg*. 2002;128:631–634.

Goldman R, Fristoe M. *Goldman-Fristoe Test of Articulation*. Circle Pines, MN: American Guidance Service; 1986.

Haapanen ML, Rantala S. Correlation between the age at repair and speech outcome in patients with isolated cleft palate. *Scand J Plast Reconstr Surg Hand Surg*. 1992;26:71–78.

Hardin-Jones MH, Jones DL. Speech production of preschoolers with cleft palate. *Cleft Palate Craniofac J*. 2005;42:7–13.

Holdsworth WG. Early treatment of cleft lip and cleft palate. *Br Med J*. 1954;4857:304–308.

Huebener DV, Marsh JL. Management of cleft lip and palate: the first 18 months. Presented at the 56th Annual Meeting of the American Cleft Palate-Craniofacial Association; April 1997; New Orleans, LA.

International Phonetic Association. *A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press; 1999.

Jolleys A. A review of the results of operations on cleft palates with reference to maxillary growth and speech function. *Br J Plast Surg*. 1954;7:229–241.

Kaplan EN. Cleft palate repair at three months? *Ann Plast Surg*. 1981;7:179–190.

Kaplan I, Ben-Bassat M, Taube E, Dresner J, Nachmani A. Ten-year follow-up of simultaneous repair of cleft lip and palate in infancy. *Ann Plast Surg*. 1982;8:227–228.

Kaplan I, Dresner J, Gorodischer C, Radin L. The simultaneous repair of cleft lip and palate in early infancy. *Br J Plast Surg*. 1974;27:134–138.

Kaplan I, Taube E, Ben-Bassat M, Dresner J, Nachmani A, Rosenbaum M. Further experience in the early simultaneous repair of cleft lip and palate. *Br J Plast Surg*. 1980;33:299–300.

Kemp-Fincham SI, Kuehn DP, Trost-Cardamone JE. Speech development and the timing of primary palatoplasty. In: Bardach J, Morris HL, eds. *Multidisciplinary Management of Cleft Lip and Palate*. Philadelphia: WB Saunders; 1990:736–745.

Kirschner RE, Randall P, Wang P, Jawad AF, Duran M, Huang K, Solot C, Cohen M, LaRossa D. Cleft palate repair at 3 to 7 months of age. *Plast Reconstr Surg*. 2000;105:2127–2132.

Koberg W, Koblin I. Speech development and maxillary growth in relation to technique and timing of palatoplasty. *J Maxillofac Surg*. 1973;1:44–50.

Kuijpers-Jagtman AM, Long RE. The influence of surgery and orthopedic treatment on maxillofacial growth and maxillary arch development in patients treated for orofacial clefts. *Cleft Palate Craniofac J*. 2000;37:527–527.

Liao Y, Cole TJ, Mars M. Hard palate repair timing and facial growth in unilateral cleft lip and palate: a longitudinal study. *Cleft Palate Craniofac J*. 2006;43:547–556.

Liao Y, Mars M. Hard palate repair timing and facial growth in cleft lip and palate: a systematic review. *Cleft Palate Craniofac J*. 2006;43:563–570.

Marks LE. *Sensory Processes*. New York: Academic Press; 1974.

Marrinan EM, LaBrie RA, Mulliken JB. Velopharyngeal function in nonsyndromic cleft palate: relevance of surgical technique, age at repair, and cleft type. *Cleft Palate Craniofac J*. 1998;35:95–100.

McWilliams BJ. Cleft palate management in England. *Speech Pathol Ther*. 1960;3:3–7.

Mølsted K. Treatment outcome in cleft lip and palate: issues and perspectives. *Crit Rev Oral Biol Med*. 1999;10:225–239.

Nollet PJ, Katsaros C, Van't Hof MA, Kuijpers-Jagtman AM. Treatment outcome in unilateral cleft lip and palate evaluated with the GOSLON yardstick: a meta-analysis of 1236 patients. *Plast Reconstr Surg*. 2005;116:1255–1262.

O'Gara MM, Logemann JA. Phonetic analyses of the speech development of babies with cleft palate. *Cleft Palate J*. 1988;25:122–134.

O'Gara MM, Logemann JA, Rademaker AW. Phonetic features by babies with unilateral cleft lip and palate. *Cleft Palate Craniofac J*. 1994;31:446–451.

Oller DK. *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.

Oller DK, Delgado RE. *Logical International Phonetics Programs* [computer software] Miami, FL: Intelligent Hearing Systems; 2000.

Peet E. The Oxford technique of cleft palate repair. *Plast Reconstr Surg*. 1961;28:282–294.

Peterson-Falzone SJ. A cross-sectional analysis of speech results following palatal closure. In: Bardach J, Morris HL, eds. *Multidisciplinary Management of Cleft Lip and Palate*. Philadelphia: WB Saunders; 1990:750–757.

Peterson-Falzone SJ, Hardin-Jones MA, Karnell MP. *Cleft Palate Speech*. 3rd ed. St. Louis: Mosby; 2001.

Randall P, La Rossa DD, Fakhraee SM, Cohen MA. Cleft palate closure at 3 to 7 months of age: a preliminary report. *Plast Reconstr Surg*. 1983;71:624–627.

Rohrich RJ, Byrd HS. Optimal timing of cleft palate closure. Speech, facial growth, and hearing considerations. *Clin Plast Surg*. 1990;17:27–36.

Rohrich RJ, Gosman AA. An update on the timing of hard palate closure: a critical long-term analysis. *Plast Reconstr Surg*. 2004;113:350–352.

Rohrich RJ, Love EJ, Byrd S, Johns DF. Optimal timing of cleft palate closure. *Plast Reconstr Surg*. 2000;106:413–421.

Rohrich RJ, Rowsell AR, Johns DF, Drury MA, Grieg G, Watson KJ, Godfrey AM, Pole MD. Timing of hard palatal closure: a critical long-term analysis. *Plast Reconstr Surg*. 1996;98:236–246.

Sandberg DJ, Magee WP, Denk MJ. Neonatal cleft lip and cleft palate repair. *AORN J*. 2002;75:490–499.

Scherer NJ, D'Antonio LL. Parent questionnaire for screening early language development in children with cleft palate. *Cleft Palate Craniofac J*. 1995;32:7–13.

Schiavetti N, Metz DE, Sitler RW. Construct validity of direct magnitude estimation and interval scaling of speech intelligibility: evidence from a study of the hearing impaired. *J Speech Hear Res*. 1981;24:441–445.

Schiavetti N, Sacco P, Metz DE, Sitler RW. Direct magnitude estimation and interval scaling of stuttering severity. *J Speech Hear Res*. 1983;26:568–573.

Semb G, Shaw WC. Facial growth after different methods of surgical intervention in patients with cleft lip and palate. *Acta Odontol Scand*. 1998;56:352–355.

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psych Bull*. 1979;86:420–428.

Smit A, Hand L, Freilinger J, Bernthal JE, Bird A. The Iowa articulation norms project and its Nebraska replication. *J Speech Hear Disord*. 1990;55:779–799.

Stevens SS. *Psychophysics*. New York: John Wiley & Sons; 1975.

Stirrat GM, Farrow SC, Farndon J, Dwyer N. The challenge of evaluating surgical procedures. *Ann R Coll Surg Engl*. 1992;74:80–84.

Stoel-Gammon C. Prelinguistic vocal development: measurement and predictions. In: Ferguson CA, Menn L, Stoel-Gammon C, eds. *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press; 1992:439–456.

Stoel-Gammon C. Role of babbling and phonology in early linguistic development. In: Wetherby AM, Warren SF, Reichle J, eds. *Transitions in Prelinguistic Communication*. Baltimore, MD: Paul H. Brookes Publishing; 1998a:87–110.

Stoel-Gammon C. Sounds and words in early language acquisition: the relationship between lexical and phonological development. In: Paul R, ed. *Exploring the Speech-Language Connection*. Baltimore, MD: Paul H. Brookes Publishing; 1998b:25–52.

Trost JE. Articulatory additions to the classical description of the speech of persons with cleft palate. *Cleft Palate J*. 1981;18:193–203.

Vihman MM. Early phonological development. In: Bernthal JE, Bankson NW, eds. *Articulation and Phonological Disorders*. Boston: Allyn and Bacon; 2004:63–97.

Vihman MM. Early syllables and the construction of phonology. In: Ferguson CA, Menn L, Stoel-Gammon C, eds. *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press; 1992:393–422.

Vihman MM. Individual differences in babbling in and early speech: predicting to age three. In: Lindblom B, Zetterstrom R, eds. *Precursors of Early Speech*. New York: Stockton Press; 1986:95–113.

Wardill WEM. Cleft palate. *Br J Surg*. 1933;21:347–369.

Weismer G, Laures JS. Direct magnitude estimation of speech intelligibility in dysarthria: effects of a chosen standard. *J Speech Hear Res*. 2002;45:421–433.

Whitehill T, Chun J. Intelligibility and acceptability in speakers with cleft palate. In: Windsor F, Kelly ML, Hewlett N, eds. *Themes in Clinical Linguistics and Phonetics*. Mahwah, NJ: Lawrence Erlbaum; 2002:405–415.

Whitehill T, Lee ASY, Chun J. Direct magnitude estimation and interval scaling of hypernasality. *J Speech Hear Res*. 2002;45:80–88.

Ysunza A, Pamplona C, Mendoza M, García-Velasco M, Aguilar P, Guerrero E. Speech outcome and maxillary growth in patients with unilateral complete cleft lip/palate operated on at 6 versus 12 months of age. *Plast Reconstr Surg*. 1998;102:675–679.

Zimmerman IL, Steiner G, Pond RE. *Preschool Language Scale-3*. San Antonio, TX: The Psychological Corporation; 1992.

Zraick RI, Liss JM. A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. *J Speech Hear Res*. 2000;43:979–988.

Appendix A

Sample No. 1

I should make my house.
yes.
I'm gonna put my beds right here so people can sleep in them.
in the toilet.
it's a bug!
it's a bug!
she goes to sleep and he goes to sleep.
yep.
where she sleep?
where she sleep?
she sleeps in the what?
she can sleep right here.
wait, she looks outside.

Sample No. 2

I found a dandy flower.
yeah.
dandy flower.
what is this?
gonna get this dandy.
another flower.
another dandy flower.
let's see if I can find another one.
nope, there are only two.
what's this?
this is for pizza.
want pizza too?
okay.
here's the pizza.
where's the pizza?