
Articulation Rate and Speech-Sound Normalization Failure

Peter Flipsen Jr.
University of Tennessee,
Knoxville

Not all children with speech delay (SD) of unknown origin develop fully normal speech even with intervention. Many retain residual distortion errors into adolescence and ultimately into adulthood. The current study examined whether articulation rate distinguishes those children who retain residual errors from those who normalize. Two groups of speech-delayed children originally identified at preschool age were retested at age 9 years (the early follow-up group) and at age 12–16 years (the late follow-up group), respectively. No differences in articulation rate were observed at either test time in conversational speech between those children who continued to produce residual distortion errors (RE) compared to those children who had fully normalized speech (NSA). For the late follow-up group, children in the RE outcome group articulated speech at significantly slower rates than the children in the NSA outcome group in an embedded words task using both syllables per second and phones per second measures. Findings suggested that children with SD of unknown origin who fail to normalize may have relative speech-motor deficits and possibly deficits in language formulation skill. Alternatively, slower articulation rate in structured tasks may represent some sort of compensation for the continuing presence of speech-sound errors. Possible motivations for such compensation are discussed.

KEY WORDS: normalization, speech delay, children, rate

Many children have speech-sound production problems that cannot be linked to any obvious specific cause; such children have historically been referred to as having functional articulation disorders¹ (cf. Weiss, Gordon, & Lillywhite, 1987). Although not all of these children develop fully normal speech even with intervention (cf. Gruber, 1999; Shriberg, Gruber, & Kwiatkowski, 1994; Shriberg, Kwiatkowski, & Gruber, 1994), the reason for such normalization failures is as yet unknown. If children who fail to normalize possess some stable constitutional difference that mitigates against normalization, understanding such differences would be invaluable to both improving intervention efforts (i.e., preventing such failures) and understanding this population in general.

¹Hereafter referred to as speech delay (SD) of unknown origin. The term *delay* is used here in a very generic sense and should not be interpreted as a condition orthogonal to *disorder*. The question of whether the error pattern exhibited by these children represents simply a temporal shift in the normal pattern of acquisition or some different pattern remains unresolved (cf. Shriberg, Gruber, & Kwiatkowski, 1994).

Articulation Rate

The current study explored the issue of normalization failure by comparing those children who normalize with those who fail to do so on articulation rate (the pace at which the speech segments themselves are produced; Turner & Weismer, 1993). Articulation rate is one component of the broader metric speaking rate, which also includes pausing. Early work suggested that the bulk of the variability in speaking rate occurred during pauses (e.g., Goldman-Eisler, 1968), but more recent analyses (e.g., Miller, Grosjean, & Lomanto, 1984) have suggested that there is considerable variability in articulation rate as well.

Articulation rate data for the current study were derived from conversational speech samples. Although conversational articulation rate indexes the time spent in the physical act of production of speech segments, it likely also includes time spent in language formulation. Both Rochester (1973) and Butterworth (1980) have made the case that pauses represent time for language formulation, but it does not necessarily follow that language formulation only occurs during pauses. Butterworth suggested that formulation may take longer than articulation, and pauses may simply represent time to catch up. Levelt (1989, p. 24) has argued that, particularly in discourse, language formulation must be occurring while speech is ongoing or significant dysfluency would result. Allen (1975) would appear to agree, having suggested that speech rhythm, of which rate is an element, is “a product of both performance universals and language-specific grammatical rules” (p. 75). Thus conversational articulation rate would appear to index a combination of both speech-motor skill and language skill. Also included in the current study were articulation rate findings from words embedded in structured phrases produced at follow-up testing. Although the embedded words task was originally intended for acoustic analysis of residual distortion errors, the task conveniently also provided the possibility for additional insight into the nature of articulation rate in the target population. Articulation rate data from such structured tasks would likely include minimal influence of language formulation on rate. Comparison of findings from the embedded words task with the conversational speech task allows for the examination of the relative influence of language and speech-motor skill on articulation rate.

The Speech-Motor Deficit Hypothesis

It has long been suggested that deficits in speech-motor skill may be partially responsible for the difficulties experienced by at least some children with SD of unknown origin. As early as 1964, Jenkins and Lohr described this issue as being of “historic interest” and

cited over a dozen studies dating back to 1926 that had examined the question. Jenkins and Lohr concluded that it was not possible to draw definite conclusions at that time. Winitz (1969) conducted a similar review, though in much greater detail, and concluded that

In general the studies do not indicate that articulatory defectives are retarded on any specific measure of general motor ability. Sufficient evidence is also lacking to support the hypothesis that articulatory defectives demonstrate a general retardation in motor skills. (p. 155)

A review of the studies since Winitz (Flipsen, 1999) concluded that there is still no definitive answer to this question. Studies that treated this population as a single undifferentiated group (e.g., Dworkin & Culatta, 1985; Jordan, Hardy, & Morris, 1978; Qvarnstrom, Jaroma, & Laine, 1993) found both differences and no differences compared to typically developing children. Studies that compared subgroups of children with SD of unknown origin to each other or to typically developing children (e.g., Bradford & Dodd, 1996; Hamlet, 1985; Hetrick & Sommers, 1988; McGlone & Proffit, 1973) also revealed a mixed picture.

The hypothesis of the current study was that if deficits in speech-motor skill are responsible for some or all of the delays, such deficits might be most obviously manifest as a failure to normalize. Thus, comparison of children who normalize with those who fail to do so might be most revealing of any speech-motor deficits present.

Normalization and Speech Delay

Several investigators have attempted to identify factors associated with normalization in children with SD of unknown origin (Andersland, 1961; Cantwell & Baker, 1987; Hall & Tomblin, 1978; Irwin, Huskey, Knight, & Oltman, 1974; Shriberg, Kwiatkowski, & Gruber, 1994; Steer & Drexler, 1960). No consistent factor or factors emerge from these studies. The only factor present in even two such studies was reported by Steer and Drexler (1960) and Irwin et al. (1974), who suggested performance on certain errors (both listed /l/) as potential predictors of outcomes, but they reported conflicting findings on overall severity as a predictor.

Given the previously discussed potential impact of language formulation on articulation rate, one other factor of relevance to the current study might be the presence of concurrent language deficits. Such concurrent deficits are not an insignificant issue; an analysis by Shriberg and Austin (1998) for example, suggested that up to 20% of speech-delayed children may have concurrent receptive language disorders and up to 60% may have concurrent expressive language disorders. At least two studies have presented data on speech-sound

normalization in the presence of concurrent language deficits. Hall and Tomblin (1978) reported that 4 (22%) of the children who initially had problems with both language and speech-sound production had continuing problems with speech-sounds as adults. At the same time, only 1 (6%) child whose only initial difficulty was with speech sounds continued that problem into adulthood. Post hoc analysis by the current author indicated no significant difference in proportions between the two groups, $\chi^2(1) = 2.090, p > .05$. Cantwell and Baker (1987) reported that 87 (80%) of the children who initially had concurrent language problems continued to have speech-sound problems at age 9 years compared to 40 (51%) of those who initially only had speech-sound problems. In this case, post hoc analysis indicated that the relative proportions were significantly different, $\chi^2(1) = 16.988, p = .000$. Given the larger sample size of the Cantwell and Baker study, failure to observe significant differences in the Hall and Tomblin study may have reflected statistical power constraints. An alternate account of these mixed findings might be that language deficits may need to be specific to a particular domain of language (e.g., syntax, morphology, or semantics) in order to have an impact on speech-sound normalization. Unfortunately, neither study provided sufficient details as to the nature of the language deficits being observed to examine this latter possibility.

Articulation Rate and Normalization Failure

A previous study (Flipsen, 2002) showed significant increases in articulation rate for the children in the current study (unselected for normalization outcome) during the course of the normal developmental period (i.e., from preschool age to after at least age 9 years). This finding held when rate was measured in either syllables per second or phones per second. Analysis by Flipsen also indicated that the rate values obtained at initial and follow-up testing overall were consistent with findings from cross-sectional studies of typically developing children when measured in syllables per second (Amster, 1984; Hall, Amir, & Yairi, 1999; Haselager, Slis, & Rietveld, 1991; Kowal, O'Connell, & Sabin, 1975; Pindzola, Jenkins, & Lokken, 1989; Walker, Archibald, Cherniak, & Fish, 1992). Relative to the phones per second measure, comparisons for the initial testing data indicated slower articulation rates being produced by the current study participants; no direct comparisons were possible for the follow-up data. It was noted by Flipsen that many of the frequent speech-sound errors in the initial samples resulted in simpler syllable structures, suggesting that the rate difference in phones per second may have been an artifact of the presence of the errors. Because such errors were largely resolved by follow-up testing, Flipsen

inferred that children with SD of unknown origin may start out with slower articulation rates but eventually catch up to their typically developing peers.

It is as yet unclear if this proposed tendency to catch up to typically developing children holds true for all children in this population or if some subgroup might fail to do so. No studies could be identified that have addressed this question. In an analogous study of preschool children who stutter, Hall et al. (1999) compared articulation rate data obtained longitudinally from 8 children whose stuttering had recovered against comparable data obtained from 8 children whose stuttering persisted and from a group of 8 typically developing children. The groups did not differ significantly on articulation rate at any of the three test times reported (initial visit, 1 year follow-up, 2 year follow-up) when rate was measured in syllables per second. However, when rate was measured in phones per second, the persistent stuttering group (i.e., those who failed to normalize) articulated speech at significantly slower rates than the typically developing children at the 1 year follow-up, but not at the other two test times.

Purpose of the Current Study

The current study was intended to examine whether differences in long-term outcomes in children with SD of unknown origin are associated with differences in articulation rate. Findings of such differences (i.e., slower rates in those who fail to normalize) in both conversational speech and the embedded words task might suggest that speech-motor deficits were at least contributing to the problem in such a subgroup. Differences in conversational speech only would suggest problems of language formulation for the subgroup. Differences in the embedded words task only would suggest that the children with continuing errors were attempting to compensate for the presence of the errors (a strategy potentially too difficult to maintain in conversational speech). Differences at initial testing only (regardless of the task) might suggest a short-term situation in which a need to compensate disappeared over time. Such a finding would be consistent with the pattern observed by Flipsen (2002) and Hall et al. (1999). Differences at follow-up testing only (again regardless of task) would suggest a subtle problem that was not manifest early on but only emerged over time.

Method

The Speech Disorders Classification System (SDCS; Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997b) served as the organizational heuristic for the current study. The SDCS serves to "classify a person's speech production status throughout the lifespan" (p. 723), thus

allowing for the consistent tracking of individuals throughout the progress of their disorder. In the context of the SDCS, the normal developmental period for speech-sound acquisition is assumed to end nominally at age 9 years. All of the participants in the current study were initially recruited at preschool age as having SD of unknown origin. At follow-up testing (age 9 years for the early follow-up group; age 12–16 years for the late follow-up group) all of the children were classified using the SDCS as either RE (producing residual speech errors) or NSA (normalized speech acquisition).

Participants

The current study involved analysis of data from the same two groups of participants as in the previous study (Flipsen, 2002). Participants were part of two separate studies and had been recruited at preschool age from the Madison, Wisconsin, area. Follow-up testing for both groups occurred concurrently. All of the children in both groups received treatment from clinicians in the community for some or all of the intervening period. The interested reader should consult the previous study for additional details on recruitment and representativeness of the participant samples.

Participant characteristics are shown in Table 1. Analysis of narrow phonetic transcriptions of the conversational speech samples obtained at follow-up testing (see below) indicated that 12 (71%) of the children in the early follow-up group and 6 (17%) of the children in the late follow-up group continued to produce residual distortion errors. Specifically they met criteria for the category RE (residual errors) using the SDCS (Shriberg et al., 1997b), which means they produced at least two

errors on a phoneme target on each of two different words. Note that the SDCS category RE includes several subtypes (depending on whether omission or substitution errors are also present), but the subtypes were collapsed into a single category (RE-A) for purposes of the current study. The remaining children in each group met criteria for the category NSA (normalized speech acquisition). Sex ratios overall for both groups (twice as many boys as girls) were typical of this population (Shriberg & Kwiatkowski, 1994). Within each of the study groups there were no significant differences in sex ratio between the NSA and RE outcome groups: early follow-up group, $\chi^2(1) = 3.192, p = .074$; late follow-up group, $\chi^2(1) = 3.600, p = .058$. Although the data in both groups trended strongly toward a difference, the trends were in opposite directions (toward more boys with RE in the early follow-up group but more girls with RE in the late follow-up group). In addition, there were no significant differences between the NSA and RE outcome groups for initial age ($ps = .9579$ and $.8151$ for the early and late follow-up groups, respectively) or follow-up age ($ps = .0962$ and $.2021$).

Because Steer and Drexler (1960) and Irwin et al. (1974) had presented conflicting findings regarding severity of involvement and normalization, and in order to rule out the possibility that severity might confound the current findings, the RE and NSA outcome groups were compared on severity of involvement using a selected set of severity metrics (Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997a). Findings are shown in Table 1. Using a series of Wilcoxon–Mann–Whitney tests, no significant differences were observed on any of the metrics, including percentage of consonants correct ($ps = 1.0000$ and $.3395$ for the early and late follow-up

Table 1. Study participant characteristics.

	Early follow-up group		Late follow-up group	
	NSA ^a	RE ^b	NSA ^a	RE ^b
N	5	12	30	6
Sex distribution (male/female)	2/3	10/2	22/8	2/4
Initial age in months ^c	49 (38–60)	49.5 (35–63)	53.5 (39–71)	53 (48–59)
Follow-up age in months ^c	112 (110–117)	110 (108–118)	176 (152–201)	171 (158–187)
Initial % of consonants correct ^f	68.2 (62.3–73.3)	67.4 (43.5–79.8)	62.7 (49.0–77.0)	65.5 (59.3–73.2)
Initial intelligibility index ^{c,d}	92.1 (77.4–97.2)	89.9 (41.5–92.3)	94.4 (71.3–99.5)	90.3 (78.8–97.8)
Initial relative distortion index ^{c,e}	13.2 (5.6–18.7)	16.6 (1.6–44.8)	27.9 (3.1–55.7)	21.6 (8.3–41.7)

^aChildren meeting criteria for normalized speech acquisition. ^bChildren meeting criteria for residual errors. ^cCell entries are group median values (and ranges). ^dPercentage of words understood in conversation. ^ePercentage of all errors that were distortions.

groups, respectively), intelligibility index (the percentage of words understood in conversation; $ps = .2684$ and $.4199$), or relative distortion index (the proportion of all errors that were distortions; $ps = .3166$ and $.4837$).

Three additional analyses were conducted to examine the potential impact of concurrent language involvement on any findings obtained. To assess receptive language skills, data on the Peabody Picture Vocabulary Test–Revised (PPVT-R; Dunn & Dunn, 1981) were examined. All but 2 of the children (both in the late follow-up group) scored within 1 *SD* of their age-group mean on the PPVT-R. Although it could be argued that the 2 children with low PPVT-R scores might have had a “known” reason for their speech delay (i.e., a possible cognitive delay), no cognitive assessment data were available for any of the children. Conversely, their reduced scores on the PPVT-R might simply indicate that they had comorbid receptive language disorders (Shriberg & Austin, 1998). In addition, an examination of the distributions for the late follow-up group indicated that these two children were not outliers in terms of segmental accuracy, intelligibility, or articulation rate at initial testing; thus, these 2 children were retained in the current study. Overall, it appeared that deficits in receptive vocabulary would not likely confound the current findings. The potential impact of concurrent expressive language deficit skills were examined by analyzing 50 utterance samples from the conversational speech samples using assigning structural stage (ASTS) analysis (Miller, 1981, p. 31) and grammatical morpheme use stage (GMUS) analysis (Paul & Shriberg, 1982). In both cases a language delay was defined using a somewhat liberal criterion of being at least one stage behind age expectations. Using the ASTS analysis, in the early follow-up group, 2 (17%) of the RE outcome group were delayed compared to 1 (20%) of the NSA outcome group, $\chi^2(1) = 0.142$; in the late follow-up group, 5 (83%) of the children in the RE outcome group were delayed compared to 14 (47%) in the NSA outcome group, $\chi^2(1) = 2.697$. Using the GMUS analysis, for the early follow-up group, 6 (50%) in the RE outcome group were delayed compared to 3 (60%) in the NSA outcome group, $\chi^2(1) = 0.142$; in the late follow-up group, 6 (100%) in the RE outcome group were delayed compared to 21 (70%) in the NSA outcome group, $\chi^2(1) = 2.400$. None of these comparisons were significant (all $ps > .05$). Findings did not change for either measure when a more conservative criterion for delay of at least two stages behind was applied to the data. Overall, regardless of the measure, the presence of language impairment (whether receptive or expressive) did not appear to be significantly associated with normalization outcomes in either study group and would not likely have confounded the findings herein.

Assessment

Both initial and follow-up testing included a large number of tasks representing several domains. As with the previous study (Flipsen, 2002), data for the current study were based on analysis of the conversational speech samples obtained at both test times using similar procedures (Shriberg, 1986). The interested reader is also referred to the previous study for details on how the conversational speech samples were obtained.

Embedded Words

In addition to the conversational speech samples, articulation rate measures at follow-up testing were also derived from productions of words embedded in structured phrases (hereafter the embedded words task) obtained as part of a larger study of the acoustics of residual speech errors (Flipsen, Shriberg, Weismer, Karlsson, & McSweeney, 1999, 2001). Recordings for the embedded words task were made using a head-mounted microphone positioned approximately 1.5 in. from the lips and no more than 2 in. from the child’s nose, with the microphone tilted toward the nose. Each child produced 5 tokens each of 10 words containing /s/ (*assign, cosine, kicks, kiss, kits, sin, skin, soon, spin, spoon*) and 12 words containing /r/ (*bird, burg, burr, cried, crude, pried, prude, rebel* (noun), *rebel* (verb), *ride, rude, tried*). These targets were selected because /s/ and /r/ are among a small set of speech sounds (along with /l/) that persist as residual distortion errors in adults (Shriberg, 1993). All 18 of the children in the current study who retained residual distortion errors were producing errors in conversational speech on either /s/ or /r/ (see below). The embedded words task also included the words *kin* and *pin*, intended to allow comparison of singleton versus cluster productions of /s/ (the subject of another investigation). The 24 target words were arranged into five different randomized orders and the resulting list of 120 words was read by the examiner, in identical order for all participants. The child’s task was to reproduce the target word in the phrase “Say _____ again.”

Measurement Procedures Conversational Speech

Analysis of articulation rate on all 106 conversational speech samples (53 initial and 53 follow-up) was conducted by the current author using digitized versions of the samples created using the Record utility of the software program CSpeech (Milenkovic, 1996). Samples were digitized using a sampling rate of 22 KHz, 15 bits of quantization, 72 dB of dynamic resolution, and low-pass filtering at 9.8 KHz. Analyses were carried out on all of the follow-up samples first, in the same order in which follow-up testing had been conducted (i.e., an order that

was neutral to outcome status). The same order was used for analysis of the initial samples.

Several investigators (e.g., Haselager et al., 1991; Malecot, Johnston, & Kizziar, 1972; Miller et al., 1984; Walker et al., 1992) have suggested that in order to minimize the influence of language formulation on articulation rate, the most appropriate measurement unit is the stretch of speech bound by pauses. Malecot et al. (1972) referred to this unit as an “utterance,” both Miller et al. (1984) and Walker et al. (1992) labeled it a “run,” and Haselager et al. (1991) called it a “phonetic utterance.” For purposes of the current study, this same unit was termed the “phonetic phrase” (after Allen, 1973). This label emphasizes an attempt to capture something closer to output behavior only (hence “phonetic”), and it avoids confusion with the more conventional meaning of the term “utterance” (hence “phrase”). Haselager et al. noted that this unit eliminates the influence of pause time on rate. Although simple subtraction of the time spent in pausing from the utterance durations might accomplish this same end, using linguistic boundaries to define the units of measurement means that there could still be considerable influence of language formulation on rate. The phonetic phrase is somewhat neutral relative to linguistic boundaries because not all pauses occur at such points (Henderson, Goldman-Eisler, & Skarbek, 1966), and thus the influence of language formulation is reduced.

For purposes of delimiting phonetic phrases, pauses were identified using wideband (500 Hz) spectrograms and initially defined as silent intervals of 250 ms or longer (Miller et al., 1984; O’Connell & Kowal, 1972; Walker et al., 1992). However, for cases in which potential pauses were bounded by two stop consonants, the two silent intervals (i.e., closures) might combine to exceed 250 ms. Therefore a conservative position was adopted for such cases and a modified criterion of 400 ms was used. Segmentation of phonetic phrases was not limited by utterance parsing. Especially in the follow-up samples, phonetic phrases did not always end at utterance boundaries unless pauses were observed. Such carryovers across utterance boundaries were rare in the initial samples.

Phonetic phrases beginning or ending with unintelligible syllables were excluded because of concerns about measurement precision. In addition, phonetic phrases were also excluded if (a) they consisted of single words or frank imitations of the examiner, (b) they were produced during obviously excited states or using a play register, (c) they contained extraneous noise or simultaneous talk by both participants, or (d) there was insufficient energy present to reliably identify the initial or final speech segment from the spectrogram. Phonetic phrases containing obvious dysfluencies (e.g., sound repetitions, prolongations) were also excluded; those containing normal

nonfluencies (e.g., whole-word repetitions or interjections) were retained. Previous analysis (Flipsen, 1999) suggested that inclusion of nonfluent phonetic phrases did not result in statistically significant differences in articulation rate in either the initial or the follow-up samples.

In an attempt to equate the size of the samples across speakers, a subsample of up to 30 phonetic phrases was then selected for each child, with phrases occurring earlier in the sample chosen before phrases occurring later in the sample. The 30 phrases were also selected to retain the overall distribution of phrase lengths available for each child to avoid possible bias based on particular phrase lengths. In cases where 30 or fewer phrases were available (9 [17%] of the initial samples; 18 [34%] of the follow-up samples), all usable phrases were included.

The precise beginning and ending of each useable phonetic phrase were identified from the spectrograms relative to the first segment in the transcript. Phrases were judged to begin at the onset of F1 energy for vowels and resonant consonants, the onset of broadband noise for fricative consonants, and the onset of the burst release for stop and affricate consonants. The reverse of these criteria were used to establish the ending of each phrase. Durations of the phonetic phrases were then recorded. Syllable counts were obtained manually from the phonetic transcriptions by the current author. A syllable was defined as any vowel, diphthong, or syllabic consonant in the child’s production; vowelized consonants (e.g., [beo] for [ber]), vowel onglides, or vowel offglides were not counted as syllables. Unintelligible syllables (where they did not interfere with identifying the beginning or ending of a phonetic phrase) were included in the counts because listeners can reliably identify syllables in unintelligible strings (Shriberg, 1986). Phone counts were obtained later from the phonetic transcriptions by a trained research assistant. Phone counts could only be obtained from phrases that were fully intelligible, resulting in a slight reduction in the number of phonetic phrases used for the calculations involving phones. Reductions occurred for 16 (94%) of the children in the early follow-up group and 22 (61%) of the children in the late follow-up group at initial testing and 7 (41%) of the children in the early follow-up group and 12 (33%) of the children in the late follow-up group at follow-up testing (in no case was fewer than 15 phrases available).

Conversational articulation rate includes many sources of variability not seen in more structured tasks (e.g., familiarity with the topic, partner responsiveness). One such source of variability in the current study (phonetic phrase length) could be controlled for directly. Note that at the group level (i.e., RE vs. NSA for both the early and late follow-up groups) there were no significant differences in length (Flipsen, 1999). The single

most appropriate unit for controlling length is unclear, however. Kent (1983) has suggested that children may be more segmental than adults in their approach to speech production. Kent reviewed findings from studies that suggest children use less frequent anticipatory coarticulation, much longer individual segments, and less articulatory undershoot in rapid speech. However, other researchers such as Nittrouer, Studdert-Kennedy, and Neely (1996) reported findings that lead to the opposite conclusion. Nittrouer et al. presented data “showing that the intrasyllabic gestures in fricative–vowel (FV) syllables were “co-produced” to a greater extent in children’s than in adults’ utterances” (p. 388). Resolving the debate between these two positions is beyond the scope of the current article, but in order to account for both possibilities, control for length was carried out in units of both syllables and phones. To control for syllable length, the approach of Haselager et al. (1991) was utilized. Phrases were grouped into short (2–4 syllables), medium (5–7 syllables), and long (8 or more syllables) categories. For long phrases, preference was given to phonetic phrases of 8–10 syllables in length. Except as noted below, within each length category a subset of the first five phrases was selected, including at least three different lengths (if available) and optimally no more than two phrases of any one length, never more than three. For the initial conversational samples, there were sufficient phrases of both short and medium length available to allow for comparisons. All 53 of the short and medium length samples included at least three phrases. For the follow-up samples, only long phonetic phrases (limited to 8–12 syllables) yielded enough data for analysis. For 52 (98%) of the children, at least three long phrases were available. Lacking precedent studies on phone length control, an examination of the distribution of phone lengths in the samples was conducted and suggested groupings into short phrases (6–10 phones), medium phrases (11–15 phones), and long phrases (16–22 phones). For the long phrases, preference was given to phonetic phrases of 16–20 phones. For the initial samples, all 53 of the short length samples and 52 (98%) of the medium length samples included at least three phrases. Note that the 1 child with insufficient medium length phrases was not the same child who had been excluded from the analysis of syllable length controlled long phrases. For the follow-up samples, again only long phrases yielded sufficient data for analysis; all 53 samples included at least three phrases.

Embedded Words

Rate measurements for the embedded words task were made from digitized versions of the samples created using the same procedures used for the conversational speech measurements. Productions of the incorrect target word were excluded, but productions

containing errors on the target sound were retained. Productions containing obvious dysfluencies or intraphrase pauses (>250 ms) were also excluded. For words containing /r/, tokens were also excluded if either of F2 or F3 for /r/ could not be reliably identified and tracked throughout the entire transition from the preceding vowel or consonant to the following vowel or consonant. Because the original intent of the task was to examine the acoustics of either /s/ or /r/ in the target words, only the target word was stored for analysis. Rate measurements for the current study were therefore derived from durations of the target words only (i.e., not the entire phrase; see Flipsen, Tjaden, Weismer, & Karlsson, 1996). Word durations in milliseconds were converted to articulation rate in syllables per second by dividing the obtained duration values into 1,000 and 2,000 for the one- and two-syllable words, respectively. Rates in phones per second were generated using the same procedures with phone counts being derived from the targets. Rates were then averaged across each child for all available tokens of each word type (i.e., /s/ words, /r/ words, *kin* / *pin*). Children were to be represented by mean values; thus, words were not included in the analysis unless at least 3 tokens were available. Children were not included in the analysis for a particular word type unless at least 65% of the words were available. For 1 child, technical problems with the tape precluded analysis. The net result was that data for 52 (98%) of the children were used for the /s/ analysis, 46 (87%) for the /r/ analysis, and 50 (94%) for the *kin* / *pin* analysis.

Measurement Reliability

Because assignment to outcome categories (i.e., RE vs. NSA) was based on analysis of phonetic transcriptions from the follow-up samples, reliability of transcription was considered important to the current analysis. At least 12 months after the initial analysis, transcriptions were repeated (to the 100 first-occurrence words criterion) on 12 (9.8%) of the samples chosen randomly from the larger study (Flipsen et al., 1999, 2001). Point-to-point intrajudge agreement for consonants was 96.9% for broad transcription and 90.4% for narrow transcription. Agreement for vowels was 90.4% for broad transcription and 82.1% for narrow transcription. Interjudge reliability was not estimated because the transcriptions of the follow-up samples had been completed by a single transcriber (note that interjudge reliability is documented for this transcriber on other similar samples; see data for “Jane” in McSweeney & Shriberg, 1995).

Reliability of the articulation rate measures on the conversational speech samples was assessed using re-measurements from 12 (11.3%) of the samples made approximately 1 month after the original measurements.

Six samples were randomly chosen from each of the 53 initial and the 53 follow-up samples. Re-measurements were made by the current author on the entire set of previously digitized utterances for each of the 12 samples selected. Mean point-to-point agreement on the location of pauses was 93.4%. Test–retest correlations (Pearson’s) on the duration measurements ranged from .96 to .99 across the 12 samples. Relative to syllable counts, exact matches were obtained on 332 (91.4%) of the recounted phrases. Reliability of the phone counts was established using a separate randomly chosen set of 12 (11.3%) of the samples (6 [11%] initial and 6 [11%] follow-up). Exact matches on phone counts were obtained on 478 (98.9%) of the recounted phrases.

Reliability for the word duration measures in the embedded words task was examined using a sample of data from 12 (9.8%) of the children from the larger study (representing 6 randomly chosen children originally tested by each of the two assistants). Re-measurements were made 4 years after the original measurements. Because one assistant was no longer available, all re-measurements were made by a single assistant. Target word durations were re-measured on the first two (40%) tokens produced by each child, resulting in a sample of 562 of 12,382 (4.5%) tokens. Findings are reported in Table 2.

Comparison Groups

The comparisons for the current study were based on SDCS analysis derived from the follow-up conversational speech samples. The net result of the SDCS analysis was that in the early follow-up group, 3 children produced sufficient errors to qualify as RE on /s/, 5 on /r/, 2 on both /s/ and /r/, and 2 on both /l/ and /r/. In the late follow-up group, 3 children qualified as RE based on sufficient errors on /s/, 2 children on /r/, and 1 on both /s/ and /r/. Thus, for the early follow-up group the comparisons involved 12 children classified as RE versus 5 children classified as NSA and for the late follow-up group the comparisons involved 6 children versus 30 children (RE and NSA, respectively).

Statistical Analysis

In the current study, comparisons between the NSA and RE outcome groups within each study group were made using nonparametric Wilcoxon–Mann–Whitney tests, owing to small cell sizes and uneven group sizes. A Type I error rate of .05 was used for each of the comparisons rather than a smaller familywise (i.e., Bonferroni-corrected) value because possible Type II errors (i.e., missing relevant differences) were considered to be of equal concern to finding erroneous differences in this exploratory study. Avoidance of Bonferroni correction is not unheard of in exploratory studies. Lahey and Edwards (1995), for example, used an even higher probability value ($p < .10$) as the significance level in their examination of several factors potentially associated with specific language impairment.

Flipsen (2002) noted that articulation rate increases significantly with age in both typically developing children and in children with SD of unknown origin. To account for this relationship, a series of regressions was carried out with age as a predictor of articulation rate. Although only 2 (6%) of the 36 regressions were significant ($p < .05$), the resulting residuals provided an indirect basis to control for any extraneous influence of age on rate. The residuals then served as the dependent measures for the nonparametric analyses used to compare the groups on both speaking tasks.

Results²

Conversational Speech

Findings for the group comparisons from the conversational speech samples (controlled by syllable length) are shown in Table 3. None of the comparisons at either test time using either rate unit was significant (all $ps > .05$) regardless of whether length was measured

²Individual child data are archived for the interested reader in a technical report at the Phonology Project Web site (<http://www.waisman.wisc.edu/phonology>).

Table 2. Reliability findings for word durations from the embedded words task.^a

Type of agreement	Absolute mean test–retest differences								Test–retest correlation coefficients ^b			
	Overall	Milliseconds			% of original values				Overall	Test–retest correlation coefficients ^b		
		/s/ words	/r/ words	kin/pin	Overall	/s/ words	/r/ words	kin/pin		/s/ words	/r/ words	kin/pin
Intrajudge ^c	7.5	5.5	9.7	4.1	2.5	1.5	3.5	2.0	.99	.98	.98	.99
Interjudge	7.6	7.5	7.3	9.7	2.5	2.0	2.6	4.2	.99	.99	.98	.86

^aBased on a sample of 12 of 122 children from the larger acoustics project. ^bSpearman ρ correlations. ^cBased on a single assistant only (see text).

Table 3. Comparisons of the SDCS outcome groups on articulation rate for the syllable length controlled samples at initial and follow-up testing.^a

SDCS outcome group	Initial testing				Follow-up testing	
	Short phrases (2–4 syllables)		Medium phrases (5–7 syllables)		Long phrases (8–12 syllables)	
	Early follow-up group	Late follow-up group	Early follow-up group	Late follow-up group	Early follow-up group	Late follow-up group
Syllables per second						
NSA ^b	3.11 (2.81–4.32)	3.16 (2.87–3.81)	3.39 (2.98–3.78)	3.50 (3.13–3.94)	4.28 (3.43–5.22)	5.41 (4.65–5.80)
RE ^c	3.05 (2.74–3.20)	3.31 (2.53–3.95)	3.57 (3.28–3.84)	3.42 (2.93–4.23)	4.21 (4.07–5.09)	4.56 (4.19–5.10)
<i>p</i> ^d	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>
Phones per second						
NSA ^b	7.71 (6.00–8.33)	7.78 (6.64–8.52)	7.50 (6.61–8.09)	7.44 (6.73–9.47)	11.16 (8.59–12.42)	12.78 (11.27–13.88)
RE ^c	6.89 (5.85–8.11)	7.65 (5.92–9.09)	8.20 (6.43–9.50)	7.85 (6.82–9.17)	10.62 (9.62–11.89)	11.25 (9.89–12.20)
<i>p</i> ^d	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>

Note. *ns* = not significant ($p > .05$).

^aCell entries are group medians (interquartile ranges are in parentheses). ^bChildren meeting criteria for normalized speech acquisition. ^cChildren meeting criteria for residual errors. ^dAnalysis based on Wilcoxon–Mann–Whitney tests using residuals from Rate \times Age regressions.

in syllables per second or in phones per second. Articulation rate in conversational speech did not appear to be differentially associated with either outcome group. Findings for the phone-length controlled samples are shown in Table 4 and again revealed no significant differences on any of the comparisons.

Embedded Words

Results from the articulation rate comparisons from the embedded words task at follow-up testing are shown

in Table 5. None of the comparisons for the early follow-up group was significant (all $ps > .05$). For the late follow-up group, however, articulation rates were significantly slower (all $ps < .05$) for those who met criterion for RE across all three word types using both syllables per second and phones per second measures.

The failure to include a Bonferroni correction in the current study increased the risk that the findings in Table 5 might represent Type I errors. To assess this possibility, effect sizes were calculated for the embedded words task. Lacking a widely accepted procedure

Table 4. Comparisons of the SDCS outcome groups on articulation rate for the phone length controlled samples at initial and follow-up testing.^a

SDCS outcome group	Initial testing				Follow-up testing	
	Short phrases (6–10 syllables)		Medium phrases (11–15 syllables)		Long phrases (16–22 syllables)	
	Early follow-up group	Late follow-up group	Early follow-up group	Late follow-up group	Early follow-up group	Late follow-up group
Syllables per second						
NSA ^b	3.72 (2.45–3.85)	3.10 (2.83–3.55)	3.23 (2.91–3.37)	3.34 (2.96–3.86)	4.17 (3.63–5.64)	4.94 (4.39–5.57)
RE ^c	3.36 (3.10–3.60)	3.29 (2.86–3.89)	3.35 (3.30–3.45)	3.32 (3.02–4.01)	4.08 (3.28–4.76)	5.11 (4.44–5.91)
<i>p</i> ^d	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>
Phones per second						
NSA ^b	7.76 (6.31–8.54)	7.41 (6.64–8.31)	7.39 (6.30–8.47)	7.76 (6.89–8.77)	10.40 (8.75–13.67)	12.68 (11.36–13.47)
RE ^c	7.65 (6.78–8.40)	7.45 (6.62–8.76)	7.66 (7.36–8.77)	7.95 (6.70–9.16)	10.46 (8.56–11.37)	11.93 (10.78–15.09)
<i>p</i> ^d	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>

Note. *ns* = not significant ($p > .05$).

^aCell entries are group medians (interquartile ranges are in parentheses). ^bChildren meeting criteria for normalized speech acquisition. ^cChildren meeting criteria for residual errors. ^dAnalysis based on Wilcoxon–Mann–Whitney tests using residuals from Rate \times Age regressions.

Table 5. Comparisons of the SDCS outcome groups on articulation rate from the embedded words task at follow-up testing.^a

SDCS outcome group	Early follow-up group			Late follow-up group		
	/s/ words	/r/ words	kin/pin	/s/ words	/r/ words	kin/pin
Syllables per second						
NSA ^b	3.12 (2.53–3.38)	3.56 (3.08–4.46)	3.78 (3.02–5.06)	3.14 (2.88–3.66)	4.39 (4.07–4.65)	4.30 (3.62–4.93)
RE ^c	2.85 (2.49–3.17)	3.70 (2.73–4.35)	4.21 (3.18–4.71)	2.47 (2.19–2.87)	3.41 (3.16–4.03)	3.37 (2.78–4.09)
<i>p</i> ^d	<i>ns</i>	<i>ns</i>	<i>ns</i>	.0192*	.0293*	.0494*
<i>d-np</i> ^e	0.23	0.13	0.33	0.50	0.60	0.65
Phones per second						
NSA ^b	9.92 (8.04–10.86)	11.55 (10.18–13.98)	11.35 (9.07–15.18)	10.19 (9.21–10.78)	13.92 (13.14–14.86)	12.91 (10.86–14.78)
RE ^c	9.09 (8.27–10.10)	11.97 (9.07–14.19)	12.62 (9.53–14.15)	8.08 (6.93–9.63)	10.80 (10.38–13.35)	10.11 (8.35–12.27)
<i>p</i> ^d	<i>ns</i>	<i>ns</i>	<i>ns</i>	.0242*	.0303*	.0464*
<i>d-np</i> ^e	0.47	0.16	0.29	0.94	1.21	0.74

Note. *ns* = not significant ($p > .05$).

^aCell entries are group medians (interquartile ranges are in parentheses). ^bChildren meeting criteria for normalized speech acquisition. ^cChildren meeting criteria for residual errors. ^dAnalysis based on Wilcoxon–Mann–Whitney tests using residuals from Rate \times Age regressions.

^eNonparametric effect size estimate; difference between medians divided by the pooled interquartile range.

* $p < .05$.

for calculating effect sizes with the Wilcoxon–Mann–Whitney test, a nonparametric estimate (*d-np*) was developed that was an analog to the standardized mean difference statistic (*d*). The estimates involved dividing the difference between the medians by the pooled interquartile range. Results are shown in Table 5. Values of *d-np* for the late follow-up group ranged from 0.50 to 1.21 ($M = 0.77$), indicating moderate to strong effects. This contrasts with weak to moderate effects (*d-np* = 0.13 to 0.47) obtained for the early follow-up group. This suggested that the significant results obtained in Table 5 were less likely to represent Type I errors.

Discussion

As noted previously, the findings from the current study may offer some insight into the relative influence of speech-motor and/or language formulation skill on articulation rate in this population. The differences observed in the embedded words task suggest that reduced speech-motor skill may be at least part of the problem being experienced by children with SD of unknown origin who fail to normalize. The failure to observe differences in conversational speech on the other hand is difficult to interpret in this regard. It might be expected that the additional demands of language formulation would only serve to exacerbate the differences observed in the embedded words task. It is very possible, however, that despite controlling for both phonetic phrase length and age, there was some as yet unidentified

source of variability in the conversational speech samples that was masking any differences. Variations in syntactic complexity are one possibility. Unfortunately analysis of syntactic complexity could not be carried out on the current samples because, as noted previously, phonetic phrases do not always map directly onto syntactic boundaries. The likelihood of real differences being masked in the conversational speech samples is indirectly supported by some additional data that were available. Three subtests (Oral Directions, Recalling Sentences, and Word Structure) of the Clinical Evaluation of Language Fundamentals–Revised (CELF-R; Semel, Wiig, & Secord, 1987) were administered as part of the larger study protocol at follow-up testing. The only significant difference between those who normalized and those who failed to do so was observed for the same groups that showed differences in the embedded words task; the children in the late follow-up group who failed to normalize scored significantly lower (Wilcoxon–Mann–Whitney $p = .0260$) on the Recalling Sentences subtest. Although this subtest includes both short-term memory and language formulation components, the failure to observe differences on the Oral Directions subtest suggests that short-term memory was not the critical component. Thus, language formulation skill cannot be ruled out as a factor in normalization failure for children with SD of unknown origin. Nor can it be ruled out as possibly having an impact on articulation rate in this population.

An alternate account of the slower rate findings is that the children in the late follow-up group who failed

to normalize were engaging in some form of compensation for their continuing difficulty with speech-sound production. And it is conceivable that the added demands of conversational speech (whether arising from language formulation or some other source) simply made it more difficult to engage in such compensation.

If compensation was indeed occurring, several possible motivations might be suggested. First, the reduced rate may represent a specific adjustment for the continuing difficulty with particular problematic sounds; recall that 22 of 24 of the targets in the embedded words task contained such sounds. However, the rate difference was observed in the targets *kin* and *pin*, which are not normally associated with residual distortion errors and were not in fact in error for any of these children. A second related possibility is a more general slowing of rate in the vicinity of problematic sounds. Although rate in the embedded words task was only measured on the target words, recall that the targets were produced in the context of "Say ____ again", in which the presence of the /s/ in "say" may have influenced rate for at least some of the speakers. This account is less likely, however, given the relatively high frequency of /s/ in the English lexicon; one would expect concurrent rate effects in conversational speech, which were not observed. Relative to both of these first two possible motivations, direct comparison of conversational phrases containing segmental errors against those containing no such errors was not feasible with the current samples because of the small size of the length-controlled samples (three to five phrases per speaker). A third possible motivation for compensation is that the rate reductions may have represented attempts to improve intelligibility (i.e., to assist the listener). A common response by adults to unintelligible children is a request to "slow down." Although intelligibility was not necessarily compromised in the follow-up conversational samples, rate adjustments may have happened at some intermediate point; such adjustments may already have served such a purpose and subsequently become strongly habituated. Compensation for the benefit of the listener is supported by the finding of differences for the older children (the late follow-up group) but not for the younger children (the early follow-up group); it seems likely that the older group would be more sensitive to listener needs (Bliss, 1984; Gallagher, 1977). However, as with the second motivation, the intelligibility adjustment argument is less likely because of the failure to observe rate differences in conversational speech. A fourth possible motivation may be rate adjustment or compensation reflecting some general insecurity about speech production. However, data from the student form of the Social Skills Rating System (Gresham & Elliott, 1990), completed at follow-up testing, makes such a possibility less likely.

Analysis indicated that the children in the RE outcome groups did not rate themselves significantly differently from those in the NSA outcome groups in terms of their ability to function in social situations. This was true even when analysis was limited to questions about social situations specifically involving speaking (Flipsen, 2001). A final possible motivation for reduced rates is the child taking on a therapy mode of production. Available data revealed a trend (albeit nonsignificant; $ps > .05$) for both RE outcome groups to have spent more time in therapy than the corresponding NSA outcome groups. The same testing setting (i.e., the same room) was used for both initial and follow-up testing and was typical of many speech clinic situations. Also, treatment tasks are frequently limited to levels of complexity below that of conversation; the fact that the current findings were limited to structured phrases then supports the possibility of compensation as a therapy production mode. Indeed the children in the current study were quite aware that their speech was being monitored; they were told so directly and the embedded words task was the only one in which they were required to wear the head-mounted microphone. A therapy mode of production cannot, therefore, be totally ruled out.

If compensation was occurring, it should not necessarily be taken to imply that such behavior was volitional. Findings from studies of the perturbation of normal articulatory movements (see Abbs, 1996) suggest that adjustments can be made at a rate faster than typical reaction times. Participants in such studies are rarely able to describe the nature of the changes they make. Also, the adjustments or compensations emerged at some point over a period of up to 10 years, and, thus, there may have been ample time for the adjustments to have become strongly habituated (i.e., what was once volitional may no longer be).

The failure to observe differences at follow-up in the early follow-up group raises the possibility that some of the children in that group had not yet finished acquiring speech-sound skills (i.e., they had not yet normalized) and that final membership in the RE and NSA outcome groups was not yet resolved for that group. Although these children were tested after age 9 years, which is generally considered the end of the normal developmental period for speech-sound acquisition, it has been suggested that children with SD of unknown origin may have a developmental period that is extended beyond that seen in typically developing children (Shriberg, Gruber, & Kwiatkowski, 1994). Given the previously discussed greater severity of involvement of the early follow-up group compared to the late follow-up group, it is conceivable that such extensions may only apply to children with SD of unknown origin who are more severely involved.

Of course, it may be that the division between those children with SD of unknown origin who normalize and those who fail to normalize is not a particularly helpful one relative to understanding this population. Such a division may simply reflect the relative skill of the clinicians treating these children or the adequacy of the treatment approaches being used. Unfortunately, only minimal information on treatment was available; many clinicians were involved who likely used a wide variety of intervention strategies.

Relative to the findings of Flipsen (2002), the findings from the current study suggest that the tendency for children with SD of unknown origin to catch up to their typically developing peers in terms of articulation rate may or may not represent a general trend. It is not completely clear whether such a pattern applies specifically to those children who fail to normalize their speech-sound production skills.

The current study suffered from several important limitations. First, no control group was used; comparison of data across studies can be problematic because of differences in recruitment criteria, sampling methods, and analysis procedures. Second, the samples used herein were not randomly selected but rather were convenience samples, limiting the generalizability of the findings somewhat. Recall also from Flipsen (2002) that participant loss in the early follow-up group from initial to follow-up testing was not random; thus, generalizability of the results for that group are especially tentative. Also, recall that data for 7 of the children in the current study were based on samples taken 6 or 12 months after initially being recruited; the influence of treatment during the intervening period may have affected the rate they used and thus modified the group level outcomes here. In addition, the current study subdivided the two study groups, rendering generalization even more tentative. Finally the current study used a univariate, nonparametric statistical approach without any adjustment for multiple comparisons. This raises the possibility that the current findings represent a Type I error. Effect size findings make this less likely.

Clinically, the occurrence of slowing in more structured tasks but not in conversation suggests the need to assess rate in both contexts because it confirms that speech behaviors are not necessarily stable across contexts especially for disordered speakers (see also Morrison & Shriberg, 1992). Also, if compensation for rate occurs in more structured tasks, analysis of production as well as monitoring of treatment progress would seem more appropriate in conversational speech rather than in structured tasks where compensation would be less likely (i.e., so that performance might be more representative of typical behavior).

Further study of the factors associated with normalization in this clinical population is clearly indicated. In particular, the findings obtained in the current study for the CELF-R suggest that additional examination of language processing as a factor may prove fruitful. Such studies offer the potential to more appropriately focus intervention efforts as well as to offer important insights into the nature of the problem being experienced by children with SD of unknown origin.

Acknowledgments

Portions of this study were part of a doctoral dissertation completed at the University of Wisconsin–Madison under the direction of Lawrence D. Shriberg, PhD. Many thanks to Dr. Shriberg and the following for their assistance: Chad Allen, Kate Bunton, Robin Chapman, Molly Erickson, Frederic Gruber, Stephen Handel, Raymond Kent, Judith Kuster, Joan Kwiatkowski, Marlys Macken, Jane McSweeney, Pearl Payne, Hye-Kyeong Seung, Lori Swanson, Gary Weismer, David Wilson, and Kathryn Yost. Finally, many thanks to the children (and their parents) who participated in the study.

Preparation of this article was supported in part by Grant DC00496 from the National Institute on Deafness and Other Communication Disorders.

References

- Abbs, J. H.** (1996). Mechanisms of speech motor execution and control. In N. J. Lass (Ed.), *Principles of experimental phonetics* (pp. 93–111). St. Louis, MO: Mosby.
- Allen, G. D.** (1973). Segmental timing control in speech production. *Journal of Phonetics, 1*, 219–237.
- Allen, G. D.** (1975). Speech rhythm: Its relation to performance universals and articulatory timing. *Journal of Phonetics, 3*, 75–86.
- Amster, B. J.** (1984). *The rate of speech in normal preschool children*. Unpublished doctoral dissertation, Temple University, Philadelphia.
- Andersland, P. B.** (1961). Maternal and environmental factors related to success in speech improvement training. *Journal of Speech and Hearing Research, 4*, 79–90.
- Bliss, L. S.** (1984). The development of listener-adapted communication by educable mentally impaired children. *Journal of Communication Disorders, 17*, 371–384.
- Bradford, A., & Dodd, B.** (1996). Do all speech-disordered children have motor deficits? *Clinical Linguistics and Phonetics, 10*, 77–101.
- Butterworth, B.** (1980). Evidence from pauses in speech. In B. Butterworth (Ed.), *Language production. Vol. 1: Speech and talk* (pp. 155–176). New York: Academic Press.
- Cantwell, D. P., & Baker, L.** (1987). Clinical significance of childhood communication disorders: Perspectives from a longitudinal study. *Journal of Child Neurology, 2*, 257–264.
- Dunn, L. M., & Dunn, L. M.** (1981). *Peabody Picture*

Vocabulary Test-Revised. Circle Pines, MN: American Guidance Service.

- Dworkin, J. P., & Culatta, R. A.** (1985). Oral structural and neuromuscular characteristics in children with normal and disordered articulation. *Journal of Speech and Hearing Disorders, 50*, 150–156.
- Flipsen, P., Jr.** (1999). *Articulation rate and speech-sound normalization following speech delay*. Unpublished doctoral dissertation, University of Wisconsin–Madison.
- Flipsen, P., Jr.** (2001, June). *Social skills and residual speech sound distortions following speech delay*. Poster presented at the 22nd annual Symposium on Research in Child Language Disorders, Madison, WI.
- Flipsen, P., Jr.** (2002). Longitudinal changes in articulation rate and phonetic phrase length in children with speech delay. *Journal of Speech, Language, and Hearing Research, 45*, 100–110.
- Flipsen, P., Jr., Shriberg, L., Weismer, G., Karlsson, H., & McSweeney, J.** (1999). Acoustic characteristics of /s/ in adolescents. *Journal of Speech, Language, and Hearing Research, 42*, 663–677.
- Flipsen, P., Jr., Shriberg, L. D., Weismer, G. G., Karlsson, H. B., & McSweeney, J. L.** (2001). Acoustic phenotypes for speech-genetics studies: Reference data for residual /s/ distortions. *Clinical Linguistics and Phonetics, 15*, 603–630.
- Flipsen, P., Jr., Tjaden, K., Weismer, G., & Karlsson, H. B.** (1996). *Acoustic analysis protocol* (Tech. Rep. No. 4). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison.
- Gallagher, T.** (1977). Revision behaviours in the speech of normal children developing language. *Journal of Speech and Hearing Research, 20*, 303–318.
- Goldman-Eisler, F.** (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.
- Gresham, F. M., & Elliott, S. N.** (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance Service.
- Gruber, F. A.** (1999). Probability estimates and paths to consonant normalization in children with speech delay. *Journal of Speech, Language, and Hearing Research, 42*, 448–459.
- Hall, K. D., Amir, O., & Yairi, E.** (1999). A longitudinal investigation of speaking rate in preschool children who stutter. *Journal of Speech, Language, and Hearing Research, 42*, 1367–1377.
- Hall, P. K., & Tomblin, J. B.** (1978). A follow-up study of children with articulation and language disorders. *Journal of Speech and Hearing Disorders, 43*, 227–241.
- Hamlet, S. L.** (1985). Speech adaptation: An aerodynamic study of adults with a childhood history of articulation defects. *Journal of Prosthetic Dentistry, 53*, 553–557.
- Haselager, G. J. T., Slis, I. H., & Rietveld, A. C. M.** (1991). An alternative method of studying the development of speech rate. *Clinical Linguistics and Phonetics, 5*, 53–63.
- Henderson, A., Goldman-Eisler, F., & Skarbek, A.** (1966). Temporal patterns of cognitive activity and breath control in speech. *Language and Speech, 8*, 236–242.
- Hetrick, R. D., & Sommers, R. K.** (1988). Unisensory and bisensory processing skills of children having misarticulations and normally speaking peers. *Journal of Speech and Hearing Research, 31*, 575–581.
- Irwin, J. V., Huskey, R., Knight, N., & Oltman, S.** (1974). A longitudinal study of the spontaneous remission of articulatory defects in 1665 school children in Grades 1, 2, and 3. *Acta Symbolica, 5*, 1–17.
- Jenkins, E., & Lohr, F. E.** (1964). Severe articulation disorders and motor ability. *Journal of Speech and Hearing Disorders, 29*, 286–292.
- Jordan, L. S., Hardy, J. C., & Morris, H. L.** (1978). Performance of children with good and poor articulation on tasks of tongue placement. *Journal of Speech and Hearing Research, 21*, 429–439.
- Kent, R. D.** (1983). The segmental organization of speech. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 57–87). New York: Springer-Verlag.
- Kowal, S., O'Connell, D. C., & Sabin, E. J.** (1975). Development of temporal patterning and vocal hesitations in spontaneous speech. *Journal of Psycholinguistic Research, 4*, 195–207.
- Lahey, M., & Edwards, J.** (1995). Specific language impairment: Preliminary investigation of factors associated with family history and with patterns of language performance. *Journal of Speech and Hearing Research, 38*, 643–657.
- Levelt, W. J. M.** (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Malecot, A., Johnston, R., & Kizziar, P.-A.** (1972). Syllabic rate and utterance length in French. *Phonetica, 26*, 235–251.
- McGlone, R. E., & Proffit, W. R.** (1973). Patterns of tongue contact in normal and lisping speakers. *Journal of Speech and Hearing Research, 16*, 456–473.
- McSweeney, J. L., & Shriberg, L. D.** (1995). *Segmental and suprasegmental transcription reliability* (Tech. Rep. No. 2). Phonology Project, Waisman Center on Mental Retardation and Human Development, University of Wisconsin–Madison.
- Milenkovic, P.** (1996). CSpeech Version 4 [Computer program]. Madison, WI: University of Wisconsin–Madison, Department of Electrical Engineering.
- Miller, J. F.** (1981). *Assessing language production in children. Experimental procedures*. Austin, TX: Pro-Ed.
- Miller, J. L., Grosjean, F., & Lomanto, C.** (1984). Articulation rate and its variability in spontaneous speech: A reanalysis and some implications. *Phonetica, 41*, 215–225.
- Morrison, J. A., & Shriberg, L. D.** (1992). Articulation testing versus conversational speech sampling. *Journal of Speech and Hearing Research, 35*, 259–273.
- Nittrouer, S., Studdert-Kennedy, M., & Neely, S. T.** (1996). How children learn to organize their speech gestures: Further evidence from fricative–vowel syllables. *Journal of Speech and Hearing Research, 39*, 379–389.
- O'Connell, D. C., & Kowal, S.** (1972). Cross-linguistic pause and rate phenomena in adults and adolescents. *Journal of Psycholinguistic Research, 1*, 155–164.
- Paul, R., & Shriberg, L. D.** (1982). Associations between

- phonology and syntax in speech-delayed children. *Journal of Speech and Hearing Research*, 25, 536–547.
- Pindzola, R. H., Jenkins, M. M., & Lokken, K. J.** (1989). Speaking rates of young children. *Language, Speech, and Hearing Services in Schools*, 20, 133–138.
- Qvarnstrom, M. J., Jaroma, S. M., & Laine, M. T.** (1993). Accuracy of articulatory movements of speech in a group of first-graders. *Folia Phoniatica*, 45, 214–222.
- Rochester, S. R.** (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research*, 2, 51–81.
- Semel, E., Wiig, E. H., & Secord, W.** (1987). *Clinical Evaluation of Language Fundamentals—Revised*. San Antonio, TX: The Psychological Corporation.
- Shriberg, L. D.** (1986). PEPPER: Programs to Examine Phonetic and Phonologic Evaluation Records. [Computer program]. Madison, WI: University of Wisconsin–Madison.
- Shriberg, L. D.** (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research*, 36, 105–140.
- Shriberg, L. D., & Austin, D.** (1998). Co-morbidity of speech–language disorder: Implications for a phenotype marker for speech delay. In R. Paul (Ed.), *Exploring the speech–language connection* (pp. 73–117). Baltimore, MD: Paul H. Brookes.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L.** (1997a). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, 40, 708–722.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L., & Wilson, D. L.** (1997b). The Speech Disorders Classification System (SDCS): Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, 40, 723–740.
- Shriberg, L. D., Gruber, F. A., & Kwiatkowski, J.** (1994). Developmental phonological disorders: III. Long-term speech-sound normalization. *Journal of Speech and Hearing Research*, 37, 1151–1177.
- Shriberg, L. D., & Kwiatkowski, J.** (1994). Developmental phonological disorders: I. A clinical profile. *Journal of Speech and Hearing Research*, 37, 1100–1126.
- Shriberg, L. D., Kwiatkowski, J., & Gruber, F. A.** (1994). Developmental phonological disorders: II. Short-term speech-sound normalization. *Journal of Speech and Hearing Research*, 37, 1127–1150.
- Steer, M. D., & Drexler, H. G.** (1960). Predicting later articulation ability from kindergarten tests. *Journal of Speech and Hearing Disorders*, 25, 391–397.
- Turner, G. S., & Weismer, G.** (1993). Characteristics of speaking rate in the dysarthria associated with amyotrophic lateral sclerosis. *Journal of Speech and Hearing Research*, 36, 1134–1144.
- Walker, J. F., Archibald, L. M. D., Cherniak, S. R., & Fish, V. G.** (1992). Articulation rate in 3- and 5-year-old children. *Journal of Speech and Hearing Research*, 35, 4–13.
- Weiss, C. E., Gordon, M. E., & Lillywhite, H. S.** (1987). *Clinical management of articulatory and phonologic disorders* (2nd ed.). Baltimore: Williams & Wilkins.
- Winitz, H.** (1969). *Articulatory acquisition and behavior*. New York: Appleton-Century-Crofts.

Received May 16, 2002

Accepted December 18, 2002

DOI: 10.1044/1092-4388(2003/058)

Contact author: Peter Flipsen Jr., Department of Audiology and Speech Pathology, University of Tennessee, 425 South Stadium Hall, Knoxville, TN 37996.
E-mail: pflipsen@utk.edu