

REGULAR ARTICLES

Treatment Effects on Speech Intelligibility and Length of Utterance in Children with Specific Language and Intelligibility Impairments

PAUL YODER, STEPHEN CAMARATA, & ELIZABETH GARDNER
Vanderbilt University

This purpose of this randomized group experiment was (a) to test the post-treatment (i.e., immediately after treatment) and follow-up (i.e., 8 months after the end of treatment) efficacy of a treatment designed to facilitate both sentence length and speech intelligibility (i.e., broad target recast), and (b) to explore whether pretreatment speech accuracy predicted response to treatment in children with severe phonological and expressive language impairment. The results support the conclusion that broad target recast facilitated follow-up speech intelligibility in children whose speech accuracy was relatively low prior to treatment.

Preschoolers who have grammatical impairments (e.g., lower utterance length than expected for age) and speech intelligibility impairments are at particular risk for continued language impairment. In this paper, *speech* refers to intelligibility of speech and *grammar* refers to use of proper word order, use of function words such as helping and linking verbs, and use of word endings such as possessive and plural that affect sentence length. These variables are important because children whose speech cannot be understood are less likely to elicit language-facilitating interactions from significant others (Camarata & Yoder, 2002). Reduced language-facilitating interaction from others is particularly problematic for such children because they might need more exposures to adult expansions and other language-facilitating utterances to benefit from such interactions as compared to typically developing children (Fey & Proctor-Williams, 2000).

A relatively high degree of co-occurrence of speech disorders and language disorders has been reported (e.g., Aram & Kamhi,

1982). Additionally, outcomes for children with simultaneous speech and language impairments are generally worse than for children with only one of the impairments (Aram & Hall, 1989). We refer to children with both speech and language impairments with non-verbal intelligence within normal limits as having a specific speech and language impairment (SSLI).

There has been relatively little focus on the simultaneous treatment of intelligibility and sentence length in children with SSLI. Segregating speech intelligibility and sentence length treatment might not be required and might be unnecessarily inefficient. If a treatment that addressed both sentence length and speech intelligibility goals in the same session were effective, it would be an important addition to the armament of available speech and language treatments and might improve the overall efficiency of treatment for children with SSLI.

Conversational recasts might be an effective way to facilitate sentence length and speech intelligibility within the same treat-

ment session. Two types of recasts are of interest herein: speech and sentence length. A speech recast is an adult utterance that immediately follows a child utterance, gives a neutral or positive evaluation of the meaning of the child's utterance, and is an exact or reduced imitation of the word(s) that the child attempted to say only using adult pronunciation of the attempted word(s). Information about accuracy of pronunciation (i.e., phonological information) is the only information added to the child's production. For example, if the child says, "This a wion [lion]," the speech recast might be "Yes, a lion." No new word order, word ending, or vocabulary is added to the child's utterance (Camarata, 1996). Sentence length recasts are similar to speech recasts except that new vocabulary or grammatical information is added to the preceding child utterance. For example, if the child says, "This lion," the sentence length recast might be "Yes, that is a lion" (Camarata, Nelson, & Camarata, 1994).

Speech recasts are hypothesized to provide information relevant to inducing rules about how speech sounds are combined to produce words (i.e., phonological rules) at moments when the child is most likely to process the information. Given a speech recast is temporally proximal to and retains the meaning of the child's prior utterance, it might be easier for the child to process that the adult pronunciation is different from his or her own (Nelson, 1989). It is important that speech recasts do not add new vocabulary or grammatical information to the child's preceding utterance. Many children with speech and language impairments have difficulty processing certain types of grammatical information (Leonard, McGregor, & Allen, 1992) and certain types of speech information (Tallal, Stark, & Mellits, 1985). When a particular recast adds both grammatical and phonological information to the child's utterance, it is likely that the child will only attend to one type of information. Experimental work indicates that children's grammatical development, not their phonological development, changes when the treatment is composed of recasts that add both grammatical and phonological information to the

child's utterance (Camarata et al., 1994; Fey et al., 1994). Speech recasts, therefore, are designed to add only phonological information.

Empirical support exists for the efficacy of speech recasts to facilitate the acquisition and generalization of specific speech sounds and intelligibility in children with severe intelligibility disorders who have no evidence of oral motor impairment, the majority of children with speech impairments (Schriberg, Kwiatkowski, Best, Hengst, & Terselic-Weber, 1986). Using specific speech sounds as targets, both infrequently and never-used speech sounds were acquired in intervention sessions and generalized to spontaneous speech samples with the mother or the clinician who did not treat the child. These results were maintained 9 months after treatment and showed replication across participants and behaviors (Camarata, 1993). These participants had intelligibility disorders of unknown etiology. Similarly, children with co-occurring intelligibility and sentence length disorders have demonstrated faster response to speech recasts than to imitation training as evidenced by fewer sessions until their first spontaneous and appropriate use of target speech sounds in communicative utterances with a nontraining adult (Camarata, 2004). Children with simultaneous deficits in speech, language, and cognitive domains have demonstrated increased speech accuracy and improved overall speech intelligibility when provided with sound-specific speech recasts (Koegel, Camarata, Koegel, Ben Tall, & Smith, 1998) and with speech recasts targeting overall intelligibility (Camarata, 2004; Smith & Camarata, 1999). In this latter study, the possible confound of increased familiarity with increased number of intervention sessions was controlled by transcribing the sessions in a random order.

Theoretical support for the efficacy of sentence length recasts is similar to that for speech recasts. Sentence length recasts are thought to be effective in facilitating sentence length because of their temporal proximity to the child's previous utterance and retention of the meaning of the child's previous utterance. These characteristics might aid the child in making necessary comparisons between his

utterance and the adult's utterance. Such comparisons might make the added vocabulary and grammatical information in the adult's utterance salient. Repeated exposure to many salient examples of word order or word ending targets might aid the child in inducing the underlying grammatical rules that affect sentence length (Nelson, 1989).

Sentence length recasts have been shown to be effective in facilitating early language targets that would result in longer utterances in several internally valid experiments. Nelson, Carskaddon, and Bonvillian (1973) found using sentence length recasts that added information about the verb phrase resulted in greater verb phrase complexity (e.g., adding the helping verb and the -ing word ending) and increased mean length of utterance (MLU) in children with typical development in the experimental group as compared to a randomized control group. In children with specific language disorders, Camarata et al., (1994) found that using sentence length recasts that specifically targeted particular word order or word ending rules (e.g., add the -ed to the end of verbs for past tense) resulted in generalized spontaneous use of the targeted word order or word ending rules.

The previously reported studies all have used either speech *or* sentence length recasts, but not both in the same session. Additionally, the previously reported studies used recasts for a limited set of very specific grammatical or speech-sound goals. The treatment tested in the present study differs from prior recast treatments in two ways. First, we intended to recast most of the children's utterances that afforded developmentally appropriate recasts, not just those affording recasts that model a limited set of goals. Using a broad definition of what constituted a recastable child utterance potentially allows more grammatical targets and speech sounds to be modeled. Given the broad nature of the outcomes that are of interest to parents and teachers, we selected broad outcomes for this study (i.e., MLU and speech intelligibility). It is probable that many grammatical targets and speech sounds would need to improve

for treatment effects to be demonstrated on these broad outcomes.

Second, both types of recasts were used within the same sessions. Clinicians attempted to speech recast all of the children's interpretable but poorly articulated utterances that afforded developmentally appropriate speech sounds to be modeled. Clinicians also attempted to use sentence length recasts for interpretable and relatively well articulated utterances that afforded developmentally appropriate grammatical goals to be modeled. Interventionists could ideally adjust the relative emphasis on intelligibility and sentence length on an utterance-by-utterance basis. We call this treatment broad target recasts (BTR).

Speech and language treatments that produce effects detected months after the end of treatment clearly are more important than those that produce effects seen only immediately after treatment. A recent review of language treatments for children with primary language disorders indicated, however, that such follow-ups are rare (Yoder & McDuffie, 2002). When they are present, they are usually very short term (e.g., 3 weeks). There is a need to test for more long-term effects of speech and language treatments on intelligibility and sentence length outcomes.

In this study, we distinguish *intelligibility* from *speech accuracy*. Intelligibility has been defined as the degree to which the speaker's intended message is recovered by the listener (Kent, Weismer, Kent, & Rosenbek, 1989). Speech accuracy is the extent to which the child accurately produces the speech sounds in the words he or she uses as compared to the adult version of the words. For example, a child saying "ba wo" for "ball roll" might be intelligible in that the adult understood the meaning of the child's production, but the speech accuracy of the production if measured in percent consonants correct (PCC) would be 25% (i.e., "b" correct, the final "l" in ball and roll incorrect, and "r" incorrect in roll). This example illustrates that assessments of speech accuracy will not necessarily tell us about a child's intelligibility. Studies of children with intelligibility problems indicate that the measures of speech accuracy (e.g., per-

centage consonants correct) explain only an average of 16% of the variance in speech intelligibility (e.g., percentage word attempts for which words are transcribed), even though both measures are derived from the same speech samples (Shriberg & Kwiatkowski, 1982; Shriberg et al., 1986). It is reasonable to consider intelligibility as the more socially important outcome. However, we recognize that speech accuracy and speech intelligibility are related constructs.

Initial speech accuracy might predict children's response to BTR. By predicting response to BTR, we mean that the size of the difference between BTR and the control group on post-treatment, follow-up intelligibility, or MLU is predicted to vary as a function of pretreatment speech accuracy. This prediction is a statistical interaction between pretreatment speech accuracy and treatment group predicting the post-treatment or follow-up outcome. Children who generally have poor speech accuracy will poorly articulate proportionally more of their utterances. If a child utterance is poorly articulated, but interpretable, the BTR model dictates that the interventionist should provide a speech recast rather than a sentence length recast. If frequency of speech recast affects intelligibility, then the intelligibility of the children with worse speech accuracy should benefit most from BTR. In contrast, if a child's utterance is articulated relatively well, the BTR model dictates that interventionists should use a sentence length recast. If frequency of sentence length recasts impacts MLU, then the MLUs of the children with relatively good speech accuracy should benefit most from BTR because these children will receive a higher proportion of recasts that facilitate sentence length.

The purpose of this study was to (a) test the main effects of BTR on post-treatment and follow-up MLU and intelligibility, and (b) test the statistical interaction between pretreatment speech accuracy and treatment group predicting the post-treatment and follow-up outcomes. That is, we tested whether treatment group differences on post-treatment and follow-up outcomes varied as a function of pretreatment speech accuracy.

METHOD

Participants

Fifty-two preschoolers (mean chronological age in years = 3.65; $SD = .71$) with SSLI participated in the research. The children met the following criteria establishing the presence of specific language impairment: (a) mean length of utterance (MLU) was at least 1.3 SD below chronological age expectation or a standard score of 80 or below on the expressive scale of the *Preschool Language Scale-3rd edition* (PLS-3; Zimmerman, Steiner, & Pond, 1992), (b) nonverbal IQs on the *Leiter International Performance Scale—Revised* (Leiter-R; Roid & Miller, 1997) above 80 (i.e., the minimum for the borderline IQ range + the SE for the test), and (c) passing a 25 db threshold hearing screening. Additionally, children scored no higher than a T score of 37 (i.e., corresponding to 1.3 SD below the mean t of 50, $SD = 10$) on the *Arizona Articulation Proficiency Scale* (AAPS; Fudala & Reynolds, 1986) to document speech accuracy impairments.

To increase the probability that the children would be appropriate for the language treatment being tested, the children had to have initial MLUs below 2.5, and use at least 10 different words in a 20-min language sample with an examiner. English had to be the only language spoken in the home. Children with evidence of oral motor disorders were excluded because the treatment presumably affected intelligibility problems due to causes other than oral motor disorders. Oral motor disorders were screened using the *Oral Speech Mechanism Screening Exam-Revised* (St. Louis & Ruscello, 1987).

Participants were 38 boys (73%) and 14 girls (27%) with an equal number in each treatment group. Seventy-one percent were Euro-American with 18 and 19 in the control and BTR groups, respectively. The remaining 15 were African-American (7), "other" (6), or Asian (1). There were no differences among treatments in racial composition of the sample. Parents' formal education levels were 1–2 years of college on average, but ranged from 10th grade to graduate school and did not dif-

Table 1*Means and Standard Deviations for Participant Descriptor Variables*

Pre-treatment variable	Control group (<i>n</i> = 26)		BTR group (<i>n</i> = 26)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Chronological age in months	44.3	7.6	43.2	9.6
Leiter IQ	101.8	10.0	102.8	14.9
PLS-3 receptive standard score	78.2	14.3	88.1	20.9
PLS-3 expressive standard score	68.6	14.8	74	7.5
Arizona Articulation Proficiency Scale <i>t</i> score	25.1	5.8	23.8	5.6
MLU in morphemes from language sample	1.65	.5	1.53	.4
Percentage child utterance attempts that are fully intelligible from language sample	51	21	47	13

Note. PLS-3 = Preschool Language Scale—Revised 3rd Edition; MLU = mean length of utterance; *There were no statistically significant differences between groups on any of these variables.

fer between groups. The parents' occupational status was above the national average and did not differ significantly between groups. Table 1 shows the means and *SDs* on the participants' chronological ages, nonverbal intelligence, productive and receptive language, articulation percentile rankings, MLU, and intelligibility. The average expressive standard score was 71.3 (*SD* = 14.8), and the average articulation *T* score was 24.5 (*SD* = 5.6). The AAPS *T* score has a population mean of 50 and a population standard deviation of 10. The sample mean score was, on average, more than 2.5 *SD* below age-expected levels. An average of 51% of the children's utterance attempts were fully intelligible by trained, but unfamiliar, transcribers, given up to three listenings per utterance. These data indicate that the sample had co-occurring and severe expressive language and articulation impairments.

Attrition. Only one participant did not meet the attendance criterion for BTR after enrolling in the study. Prior to enrolling, 12 participants declined to participate in the study. Seven (58%) of these had been randomly assigned to the BTR group. The most frequent reasons given for declining was too much travel and too many other things to do for these families. The remaining five families gave no reason for declining. There were no obvious differences in selection criteria be-

tween participants who declined and those that remained in the study.

Research Design

A randomized group experiment was used to test the efficacy of BTR on MLU and intelligibility at post-treatment (i.e., 6 months after study entry) and follow-up (14 months after study entry) periods. Random assignment was implemented using a computer program explicitly designed to allow equal probability of being assigned to either group. The effectiveness of random assignment was evaluated through tests of pretreatment group equivalence on 12 pretreatment variables and two nonproject treatment attendance variables. An exploratory, correlational design element was added to this randomized group experiment to test whether pretreatment speech accuracy predicted response to BTR at the post-treatment and follow-up periods.

Procedures

Overview. Upon entry into the study, children were administered the Leiter-R, PLS-3, AAPS, the oral mechanism screening, and an initial 20-min language sample. If selection criteria were met, children were administered a second 20-min language sample to provide a more stable estimate of their initial intelligibility and MLU. After pretreatment testing, children were randomly assigned to one of

two groups: BTR or control. Children in the BTR group received three 30-min treatment sessions per week for 6 months. Children in the control group were free to participate in community-based treatments but were not provided BTR. The extent to which children received outside-of-project treatment was measured to determine whether potential confounds with group assignment could account for post-treatment or follow-up differences between treatment groups. Implementation of the BTR treatment was examined in detail by coding the three treatment sessions that occurred at the 3rd, 4th, and 5th month of treatment. At post-treatment (6 months after study entry) and at follow-up (14 months after entry into the study), two 20-min language samples each were administered to measure intelligibility and MLU. The three measures of pre-treatment speech accuracy were derived from the AAPS and language samples. Outcomes were derived from the post-treatment and follow-up language samples.

Measures

Arizona Articulation Proficiency Scale (AAPS; Fudala & Reynolds, 1986). The AAPS is a standardized single-word articulation test. The child is shown a picture and asked "What is this?" If the child does not respond, the examiner provides the correct word and says, "Now you say it." In addition to using T scores (a type of standard score) to qualify and describe participants, AAPS raw scores were used to quantify elicited single-word speech accuracy that is weighted for developmental level of the target phoneme being tested. Standard scores are adjusted for chronological age; raw scores are not. As we intended to measure it, the construct of speech accuracy is referenced to adult pronunciation, and thus, for the purposes of this study, were not adjusted by children's chronological age. Raw scores are more appropriate than standard scores for our purposes of quantifying speech accuracy level. Although raw scores are not intervally scaled, the statistics used in this report do not assume the variables to be on an interval scale or ratio scale (Tabachnick

& Fidell, 1996). Internal consistency on the current sample for the raw score is .89.

Twenty-minute language samples. This procedure was administered to derive measures of intelligibility and sentence length at the three measurement periods. Additionally, the pretreatment language sample was used to derive percentage vowels correct and percentage consonants correct. These measures were derived from the average of two, 20-min language samples. The same sets of toys were used for all children and examiners were taught to use a prescribed, responsive interaction style during the language sampling sessions. To avoid boredom, separate sets of toys were counterbalanced across measurement periods.

We instituted a strict definition of a *word* to achieve reliability on the intelligibility measure. Conceptually, we transcribed any audible (whispered or vocalized) approximation of a word that is in the English dictionary and for which there was immediate nonlinguistic support, immediate conversational support, or evidence of productive use elsewhere in the session. Over- or under-extended meanings were acceptable as long as they had immediate nonlinguistic support. An acceptable approximation must be similar to the adult pronunciation in one of two ways. If the syllable structure was intact, an accurate vowel or an accurate initial consonant had to be present. In the absence of the same number of syllables as the adult pronunciation of the word, at least one accurate consonant-vowel or vowel-consonant combination had to be present.

The language samples were transcribed, indicating bound morphemes such as regular past tense (e.g., /ed/) using the format prescribed in *Systematic Analysis of Language Transcripts* (SALT; Miller & Chapman, 1993). Pilot work on the potential effect of differential transcriber familiarity indicated that familiar transcribers' MLU and intelligibility estimates were over 1 *SD* larger than unfamiliar transcribers' (Yoder & Gardner, 2004). The language samples, therefore, were transcribed by someone other than the child's clinician. Transcribers were equally unfamiliar with control and BTR participants. Using the

SALT program, we derived MLU in morphemes and the proportion of utterance attempts that were fully intelligible from pre-treatment, post-treatment, and follow-up language samples. The SALT program derives MLU from complete and fully intelligible utterances. The intelligibility measure was derived from all of the child's utterances.

After transcription was completed, the SALT file was imported into *Computerized Profiling* (CP, Long & Fey, 1995). Trained speech-language pathologists indicated the sounds that were said for each of the transcribed words. The phonology module within CP (i.e., PROPH+) was used to derive percentage vowels correct and percentage consonants correct at pretreatment. These variables were viewed as measures of pretreatment conversational speech accuracy.

Treatment Groups

Broad target recasts group (BTR). BTR sessions were conducted three times a week for 30 mins a session for 6 months. The interventionist was a woman with an undergraduate degree in psychology. She was trained to 95% compliance to the BTR model and was monitored weekly by a master's level speech-language pathologist. Therapy sessions were administered in a small therapy room using a 1:1 format. A table and chair were available in the room, but much of the interaction occurred on the floor. The therapist followed the child's attentional and play lead and asked questions about what the child was attending to and doing in an attempt to get the child to speak. When the child did speak, the interventionist determined whether the child utterance afforded a recast of a developmentally appropriate grammatical target or speech sound. Theoretically, the interventionist used sentence length recasts for well articulated utterances (i.e., few or no speech accuracy errors) and speech recasts for understood, but poorly articulated (i.e., many speech accuracy errors) utterances. The target recast rate was 4 per min. Participants were allowed to miss three treatment sessions and were asked to make up any additional sessions missed. Participants who missed treatment in excess of nine sessions

were withdrawn from the study. Only one participant that began treatment did not meet this criterion.

Control group. Children in the control group were not administered the BTR treatment. A wait-list control was not used because a major question of the study was long-term efficacy of the treatment, thus requiring a long interval between the treatment phase ending and the follow-up assessment. The participants in the control group were not prohibited from seeking and obtaining treatment outside the study.

Fidelity of treatment. To describe the BTR treatment implementation, the middle 15 mins of the videotaped treatment sessions from each of the 3rd, 4th, and 5th months of treatment were coded in detail for every child in the BTR group. The behaviors coded from these videotaped sessions are listed in Table 2.

Nonproject Treatment Description

Because families were free to enroll in community-based treatments (i.e., those offered by nonproject personnel) and because it is possible that children in the control group sought nonproject treatment more often than the BTR group, nonproject treatment could potentially relate to between-group differences in the outcomes. Although this is rarely measured in treatment studies, it was included herein because of this potential impact. Assuming that nonproject treatment can be facilitative of development and that BTR also facilitates development, greater nonproject treatment could potentially result in underestimating the BTR treatment effect size. Although control group parents were not asked to abstain from nonproject treatment, we can statistically control between-group differences on nonproject treatment involvement if there are between-group differences in the amount of nonproject treatment and such differences are related to outcomes. Therefore, we asked parents to report the number of hours their children attended speech, behavior, physical, and occupational therapy for speech and language delays in an early intervention program, daycare, clinic, or home. Because enrollment and attendance in these therapies were expected to

vary during the study, parents were asked to report participation in nonproject treatments 4 times over the course of the project (pretreatment, 6 months after study entry, 10 months after study entry, and 14 months after study entry). It should be noted that the last 2 periods occurred after the post-treatment assessments. From this information, we derived (a) the average number of therapy hours from a speech-language pathologist in all settings per month and (b) the average number of hours of any type of educational services from professionals or paraprofessionals other than speech-language pathologists. We also examined whether these two variables assessed at the last 2 measurement periods could account for differences in follow-up results.

Reliability

We estimated the interobserver reliability of the variables that were derived from transcriptions. The interobserver reliability estimate was the intraclass correlation coefficient (Suen & Ary, 1989). For all coded variables, including those derived from phonemic transcription, reliability was estimated from the videotapes of the samples. The same orthographic transcription should not be used to estimate reliability for percentage vowels and consonants correct because doing so bypasses estimating measurement error due to different orthographic transcriptions. We sampled between 17% (i.e., 8 for percentage vowels correct at pretreatment) and 39% (i.e., 20 for MLU and intelligibility at pretreatment) of the 52 participants for reliability estimates. Samples used for reliability estimation were selected randomly and reliability transcription and coding was conducted independently from the primary transcription and coding.

At pretreatment, the intraclass correlation coefficients were .98, .51, .43, and .91 for MLU, percentage utterance attempts that were fully intelligible, percentage vowels correct, and percentage consonants correct, respectively. At post-treatment, intraclass correlation coefficients were .77, and .74 for MLU and percentage utterance attempts that were fully intelligible, respectively. At follow-up, intraclass correlation coefficients were .97, and .89

for MLU and percentage utterance attempts that were fully intelligible, respectively. Intraclass correlation coefficients were between .71 (rate of child recastable utterances) and .97 ($M = .92$; $SD = .08$) for fidelity of treatment variables. Decreased reliability results in a reduced probability for detecting a significant difference in the analyses. Significant findings occurred despite reliability under .70 at Time 1.

RESULTS

Preliminary Analyses

Prior to addressing the research questions, we conducted preliminary analyses to (a) examine pretreatment differences between treatment groups, (b) test group differences on *nonproject* treatment attendance, (c) describe the amount of recasts provided to document *dosage* of treatment, and (d) document the amount of growth children experienced regardless of treatment group.

Test of between-group differences on select pretreatment variables. A test of between-group differences on select pretreatment variables was completed to determine whether random assignment yielded equivalent groups at the pretreatment period. Twelve pretreatment variables were tested for group differences to determine whether random assignment was effective in creating equivalent groups prior to treatment onset. There were no statistically significant differences on any of these 12 variables, indicating that random assignment was successful at creating equivalent groups at the pretreatment period.

Test of between-group differences on nonproject treatment participation. We tested the significance of between group differences on the average number of hours of speech-language therapy per month and the average number of therapeutic hours of other professionals and paraprofessionals per month across all four measurement periods. There were both variance, $F(2, 50) = 14.5$; $p = .0001$, and mean differences, $d = -.63$; $t(25.3) = 2.29$; $p = .03$, favoring the control group on the number of hours with other professionals and paraprofessionals. Considering both BTR and nonproject treatments, there

Table 2*Description of BTR Implementation across the 3rd, 4th, and 5th Months of Treatment*

BTR implementation strategies	<i>M</i>	Median	<i>SD</i>	Min	Max
Rate of adult speech recasts	2.6	2.8	1.4	.46	5.1
Rate of adult sentence length recasts	2.1	2.2	1.0	.62	4.5
Proportion of child recastable utterances adult speech recasts	.47	.46	.22	.12	.84
Proportion of child recastable utterances adult sentence length recasts	.35	.29	.20	.08	.78
Proportion of adult utterances that were speech recasts	.15	.14	.08	.03	.40
Proportion of adult utterances that were sentence length recasts	.17	.16	.10	.04	.56
Proportion of adult recasts that were sentence length recasts	.47	.44	.22	.15	.87

Note. BTR = broad target recast.

was a statistically significant difference in the total amount of treatment favoring the control group, $t(28.5) = -3.18, p = .004$. This between-group difference in total treatment was due to nonproject treatment provided by other professionals and paraprofessionals (i.e., not a speech-language pathologist). Additionally, we tested whether the treatment groups were different on the two nonproject treatment attendance variables at 10 and 14 months after study entry. The 10-month measurement period was between the post-treatment and follow-up measurement periods for MLU and intelligibility. The 14-month measurement period was the follow-up measurement period of MLU and intelligibility. There was no statistical evidence of group differences in nonproject treatment attendance at these periods (all p values $> .16$).

We examined whether the group difference in average amount of nonspeech/language treatment across all four measurement periods could account for the results of the tests of the research questions. Attendance to nonproject treatment was not associated with outcomes at either post-treatment or follow-up periods. Additionally, nonproject treatment attendance variables measured for the period between post-treatment and 8 months after treatment ended were not associated with MLU or intelligibility at the follow-up period. Nonproject treatment attendance, therefore, cannot account for differences between groups or interactions between group and pretreatment speech accuracy.

Description of the BTR treatment imple-

mentation. Children assigned to the BTR group attended an average of 74 treatment sessions ($SD = 5.6$). Individual variability in attendance to BTR did not account for variance in post-treatment or follow-up outcomes. Table 2 shows the central tendency and variability of the variables measured from the BTR sessions. These fidelity of treatment data document that the treatment was implemented as intended. The target rate of recasts (i.e., 4/min) was slightly surpassed, and there was much variability in the rate of recasts. This occurred in part because the rate of recastable utterances, which serve as the antecedents to staff-delivered recasts, varied widely. On average, 82% of the children's recastable utterances were recasted. The proportion of staff-implemented recasts that were sentence length recasts was positively related to the children's intelligibility at post-treatment ($r = .45; p = .05$), suggesting that more intelligible children received proportionally more sentence length and proportionally fewer speech recasts.

Growth in MLU and intelligibility, regardless of treatment group. The main effect of time was tested using a repeated measures ANOVA with time as a 3-level within-subjects factor (pretreatment, post-treatment, follow-up). We used the Greenhouse-Geisser adjustment to the degrees of freedom to correct for potential violations to the sphericity assumption. There was a strong effect of time on MLU, $F(1.43, 50) = 67.18; p < .001; \eta^2 = .62$, and a strong effect of time on intelligibility, $F(1.77, 50) = 10.89; p < .001; \eta^2 = .24$. These data indicated significant and large

growth in MLU and intelligibility across pre-treatment through follow-up.

Tests of the Research Questions

Tests of the main effect of BTR. Using four independent *t* tests, we tested whether there were BTR versus control group mean differences at the post-treatment and follow-up periods on MLU and intelligibility. There were no statistically significant differences between groups on either outcome at either measurement period (all *p* values > .38).

Statistical interactions between pretreatment speech accuracy and treatment predicting post-treatment and follow-up outcomes. As an exploratory analysis, we examined whether pretreatment speech accuracy (an attribute variable measured on a continuous scale) interacted with the treatment group (BTR vs. no BTR) in predicting individual differences in MLU or intelligibility at post-treatment (i.e., short-term) and follow-up (i.e., long-term) measurement periods. A multiple regression analysis was chosen for its ability to determine proportion of the variance in a criterion variable related to continuous (speech accuracy) and categorical (treatment group) variables and their interaction. Using a continuous form of pretreatment speech accuracy variable retained maximum information about individual differences in initial speech accuracy and maximized statistical power (Cohen, 1983). Four multiple regressions in which the criterion variable was either MLU or intelligibility at either the post-treatment or follow-up period were tested. The same three predictor variables were included in the regression model for each of the four analyses: (a) the pretreatment variable, (b) the treatment group, and (c) the product term for (a) and (b) (Aiken & West, 1991). The pretreatment variable is grand mean centered to reduce multicollinearity between the pretreatment variable and the product term (Aiken & West, 1991). The treatment group variable is dummy coded (e.g., 0, 1) to aid interpretation (Aiken & West, 1991). The test of the significance of the product term tests whether the magnitude of the between-group difference varies as a function of the pretreatment vari-

able (i.e., a statistical interaction between group and the pretreatment variable). Statistically significant interactions are interpreted using the Johnson-Neyman technique, which indicates the values on the pretreatment variable below which or above which the treatment groups differ on the dependent variable (Aiken & West, 1991). Only interactions with at least two participants in each treatment group were interpreted.

All statistical assumptions were met for interactions between pretreatment variable and group predicting post-treatment and follow-up outcomes (i.e., linearity, homogeneity of residuals, no undue influence, and low multicollinearity). Table 3 shows the details of the statistically significant interactions. BTR facilitated short-term and long-term generalized MLU in children who began treatment producing less than 49% or 50% consonants correct, respectively. Similarly, BTR facilitated long-term generalized intelligibility in children who began treatment with less than a raw score of 46 on the AAPS. These effect sizes for the entire group are weak to moderate. The total number of children in the region of significance in which BTR is superior has 9, 10, and 11 children for the post-treatment MLU, follow-up MLU, and follow-up intelligibility, respectively. Regressions are based on the entire sample of 52 children and the variables analyzed are continuous rather than dichotomous. Figure 1 shows the statistical interaction of pretreatment AAPS raw score and group predicting follow-up intelligibility. The figure shows the two regression lines representing the association between grand-mean centered AAPS raw score at pre-treatment and follow-up intelligibility for the two groups. A regression line is a line that best fits the intelligibility scores at each of the pretreatment AAPS raw score values. It is similar to a *running mean* intelligibility score at each level of the pretreatment AAPS raw score. The vertical line at -10.25 indicates that the group's predicted intelligibility scores are statistically different in children with pretreatment grand-mean centered AAPS raw scores of -10.25 and below. The mean for pre-treatment AAPS is 55.89. The uncentered pretreatment AAPS

Table 3
R², t, and p values, Regions of Significance, and Effect Sizes for Affected Subgroups in Statistically Significant Interactions

Outcome	Pretreatment variable interacting with treatment group	R ² change for interaction		t for interaction		p for interaction		Uncentered value on pretreatment variable below which the treatment groups differ
		term	term	term	term	term	term	
MLU at post-treatment	Percent consonants correct	.12	-2.06	.01	49			
MLU at follow-up	Percent consonants correct	.09	-2.3	.03	50.86			
Intelligibility at follow-up	Raw score on the AAPS	.08	-2.02	.03	45.63			

Note. MLU = mean length of utterance; AAPS = Arizona Articulation Proficiency Scale.

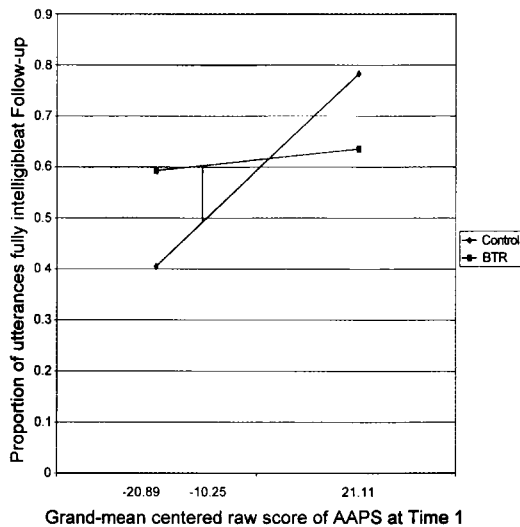


Figure 1.
Regression lines for the statistical interaction between pretreatment AAPS raw score and group predicting follow-up intelligibility.

raw score at which average groups' intelligibility scores were statistically different was 45 and under (i.e., 55.89 - 10.25 = 45.64).

To test for potential alternative explanations for the statistical interactions, we computed the effect size for the difference between control and BTR groups for all pretreatment variables within the children in the lower region of significance. Only six variables had effect sizes over a Cohen's *d* of .50. None of these pretreatment variables correlated significantly with the post-treatment MLU or either follow-up outcome. Pretreatment differences between groups within the children in the region of significance, therefore, could not explain the significant statistical interactions.

DISCUSSION

This study was conducted to determine if BTR was effective in facilitating MLU and intelligibility 8 months after the end of treatment and to determine whether pretreatment level of speech accuracy was a predictor of long-term response to BTR in a sample of preschoolers with severe expressive language and speech intelligibility impairments but age-appropriate nonverbal intelligence and normal

hearing. Children with both speech and language impairments are at particular risk of future language deficits and subsequent academic risk because poor speech accuracy is thought to reduce the frequency of potentially facilitating input from others.

The results show no evidence that the BTR treatment was effective for children across the entire range of speech accuracy represented in the present study. Instead, the BTR treatment effects were dependent on the pretreatment speech accuracy levels of children. Only children with relatively low pretreatment speech accuracy benefited from the BTR treatment. This conditional treatment effect occurred for both MLU (at the post-treatment and follow-up periods) and intelligibility (at the follow-up period).

Absence of BTR Main Effects

Several factors are potential explanations for why there was no main effect for BTR. This study was not strictly a replication study. Most previous recast studies used target-specific recasts and either sentence length or speech recasts, but not both. We used broad target recasts and combined both types of recast in one session. Additionally, the control group had greater non-SLP treatment attendance than the BTR group. Either of these explanations could account for the absence of main effects for BTR. Finally, the absent main effect of BTR could be a function of the children's need for more targeted stimulation of a limited set of particular goals.

Greater Interpretability of the Statistical Interaction on Intelligibility

Although there were three statistically significant interactions, there are two reasons why we consider the BTR effect on the intelligibility of children with relatively low pretreatment speech accuracy the most reliable (i.e., most likely to replicate). First, the children affected by BTR on intelligibility were predictable before treatment began. Because children with initially low speech accuracy were expected to receive proportionally more speech recasts and because speech recasts are the aspect of the BTR thought to address intelligi-

bility, we expected them to benefit most from BTR. This finding was seen in the results. We did not see the predicted relationship for MLU, because contrary to the findings, we expected children with relatively high initial speech accuracy to benefit from BTR with regards to MLU outcomes. Second, the BTR treatment effect on MLU became nonsignificant after controlling for pretreatment predictors of MLU (i.e., receptive language level; Yoder, Camarata, & Gardner, in preparation). As expected, the BTR effect on intelligibility of initially low speech accuracy children remained significant after controlling for pretreatment predictors of later intelligibility (Yoder et al., in preparation).

Rationale for Interpreting a Treatment Effect on Follow-up Intelligibility When Post-treatment Effects are Not Seen

Some readers might wonder how one can interpret the conditional treatment effects of BTR on follow-up intelligibility when there is no analogous effect on intelligibility at post-treatment. This type of treatment effect has been called a *sleeping effect* because the effect is not seen until months after the treatment ends (Clarke & Clarke, 2004). Such effects can occur when small changes occur at the post-treatment period in the children that elicit from their caregivers facilitating input during the interval between post-treatment and follow-up periods. These inputs, in turn, can affect follow-up child outcomes. In the present study, it is possible that our measures of intelligibility at post-treatment were not sufficiently sensitive to detect treatment effects on some children's intelligibility, which in turn elicited speech recasts from parents in the form of confirmation requests. Other reports have found that parents of children who are very unintelligible tend to use recasts after partially intelligible utterances (Yoder, Hooshyar, Klee, & Schaffer, 1996). We were not able to test whether this mechanism accounted for the conditional BTR effect on follow-up intelligibility in the present study. Similar effects, however, were seen from a prelinguistic treatment on language development that did not become manifest until 6 and 12 months

after the treatment ended (Yoder & Warren, 2001b). This indirect effect occurred through parental linguistic mapping of children's pre-linguistic communication (Yoder & Warren, 2001a).

The Paucity of Evidence about Treatment Efficacy on Intelligibility

Even though the effect size is small to moderate ($R^2 = .09$) and the effect is conditional on initial speech accuracy, the finding that BTR has a long-term effect on intelligibility is potentially important. No other studies with this population have examined treatment effects on long-term follow-up intelligibility; therefore, we cannot compare this effect size with those in comparable studies. Additionally, almost every other study examining the efficacy of speech therapy examines speech accuracy, rather than speech intelligibility, as the outcome. The almost complete lack of evidence of treatment efficacy on intelligibility, rather than speech accuracy, is striking given the social importance of intelligibility. BTR's effect on intelligibility in children with initially low speech accuracy is even more noteworthy because it occurred 8 months after treatment ended. To our knowledge, this is the first time it has been shown that conversational intelligibility is affected by treatment 8 months after the end of treatment. Documenting this finding is the most important contribution of the current study to the literature.

The finding that BTR affected the speech intelligibility of only the most developmentally immature speech producers might contradict many reader's expectations. Speech therapists have been conducting very structured articulation therapy for many decades. A continuum of support model would prescribe using BTR to address generalization of sounds recently acquired in the clinic or to address the phonological needs of less impaired children regardless of developmental level (Camarata, 1995, 1996). The present study's findings do not support such a model in that BTR (which would be considered relatively low support and highly naturalistic in interaction style) was not effective in children with relatively high levels of pretreatment

speech accuracy, but provided support that BTR affected intelligibility in children with relatively poor pretreatment speech accuracy.

Limitations of the Study

Three primary limitations existed in the present study. First, we carefully coded and thus described three sessions per BTR participant (4% of the sessions). Coding more sessions would be more likely to provide a more stable estimate of the typical implementation of the BTR treatment. Second, we did not have information available before the study to know which raw scores defined the margins of the subgroup in which BTR would be effective. This particular subgroup could have been oversampled and improved the probability that our subsample of affected children represented the population of children affected by BTR. Instead, the variables expected to predict BTR response were kept continuous and allowed the distribution of scores to be determined by chance. As usual, the tails of the distribution on the predictors of BTR response had fewer people than the middle of the distribution. Regions of statistical significance and significance level of statistical interactions are determined by the entire range of scores, not just those in the regions of significance (Aiken & West, 1991). Although replication of the interaction is needed, results of this study should not be disregarded as unimportant simply on the basis of relatively few participants occurring in the region of significance. The number of children in the region of significance in this study is similar to studies that examine predictors of language treatment response in children with disabilities (Cole, 1995; Yoder, Kaiser, & Alpert, 1991; Yoder et al., 1995; Yoder & Warren, 1998). Third, theory allowed us to predict which children BTR would affect with regards to intelligibility, but not which children BTR would affect with regards to MLU. Finally, we did not have a norm-referenced measure of expressive language. We could not determine if any of the control group children caught up with their age-appropriate expressive language without treatment. Possibly including such "late bloomers" in the study influences

our understanding of the appropriate types of children to which the findings might generalize, but does not influence the internal validity of the study.

Strengths of the Study

The quality of this efficacy study was noteworthy for several reasons. For example, the ecological validity of the outcomes was strong. They were obtained from language samples instead of standardized tests. The former is the more frequent language use context for preschoolers. Additionally, length of utterance and intelligibility are outcomes at a broad enough level of description that parents and teachers are highly likely to notice a difference. Another strength of the study is the use of an 8-month follow-up period to determine whether treatment effects occurred many months after the end of treatment. Finally, there are three reasons why internal validity of the study was strong: (a) random assignment to treatment groups was used, (b) pretreatment equivalence on a number of variables that can affect variance in outcomes was demonstrated, and (c) nonproject treatment attendance did not account for differences between treatment groups or statistical interaction between pretreatment speech accuracy and group on outcomes. Although nonproject treatment participation can relate to treatment effects, few studies examine this factor.

Measuring nonproject treatment: A novel design element. We considered it problematic to require control group parents to refrain from all other treatment for 14 months (6 months of treatment phase + 8 months of follow-up). Because group differences can and did occur after the random assignment to treatment groups, it was important to control for such differences. We did so by (a) testing to see if the confound occurred, (b) testing to see if the difference was associated with the outcomes, and (c) statistically controlling for them, if needed. The greater treatment attendance (i.e., differences on nonproject, nonspeech/language treatment) in the control group could not account for the reported significant results for two reasons. Nonproject nonspeech/language treatment was not asso-

ciated with any outcome at any period and the control group was the group with the most total and nonproject treatment. Unless treatment negatively affected outcome, it is unlikely that the nonproject treatment attendance explained why BTR children with initially low speech accuracy had better intelligibility at follow-up.

Future studies should collect nonproject treatment attendance, as it has been positively associated with later outcomes in other studies of preschoolers with disabilities (Stone & Yoder, 2001). Future research also should develop predictive measures of the type and quality of nonproject treatment and how these might modify within-project treatment efficacy. Measuring only attendance in this study is not a statement that attendance is sufficient information. The field presently almost never measures any aspect of nonproject treatment. Attendance to such treatment is only a place to start investigating this important, but rarely studied, class of potential threats to internal validity.

Future Research

Future studies are needed to determine if BTR is effective when used in the classroom by classroom teachers who are not carefully supervised. One potentially important attribute of BTR treatment is the rate of speech recasting, which averaged 2.6/min in the present study. If classroom teachers or other interventionists are spreading the number of recasts among 4 to 5 children, the number of recasts per child per min is likely to be less and the efficacy of the recast might be compromised. Additionally, it is less likely that children with SSLI will process a recast of another child's utterance as well as a recast to their own utterance because presumably recasts are effective in part because they maintain the speaker's meaning.

Conclusion

This study was the first to identify that initial single word speech accuracy predicts which children respond to broad target recasts as a new treatment for long-term conversational intelligibility. This treatment was only effec-

tive for children with initially very low single word speech accuracy (i.e., below 46 raw score on the AAPS).

REFERENCES

- Aiken, L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.
- Aram, D., & Hall, N. (1989). Longitudinal follow-up of children with preschool communication disorders: Treatment implications. *School Psychology Review, 18*, 487–501.
- Aram, D., & Kamhi, A. (1982). Perspectives on the relationship between phonological and language disorders. *Seminars in Speech, Language, and Hearing, 3*, 3–22.
- Camarata, S. (1993). The application of naturalistic conversation training to speech production in children with speech disabilities. *Journal of Applied Behavior Analysis, 26*(2), 173–182.
- Camarata, S. (1995). A rationale for naturalistic speech intelligibility intervention. In M. E. Fey, J. Windsor, & S. F. Warren (Eds.), *Language intervention: Preschool through the elementary years* (pp. 63–84). Baltimore: Brookes.
- Camarata, S. (1996). On the importance of integrating naturalistic language, social intervention, and speech-intelligibility training. In L. K. Koegel, R. L. Koegel, & G. Dunlap (Eds.), *Positive behavioral support: Including people with difficult behavior in the community*. (pp. 333–351). Baltimore: Brookes.
- Camarata, S. (2004). *A comparison of imitative and conversational recast intervention for phonological disorders in children*. Manuscript in preparation.
- Camarata, S. M., Nelson, K. E., & Camarata, M. N. (1994). Comparison of conversational-recasting and imitative procedures for training grammatical structures in children with specific language impairment. *Journal of Speech and Hearing Research, 37*(6), 1414–1423.
- Camarata, S., & Yoder, P. (2002). Language transactions during development and intervention: Theoretical implications for developmental neuroscience. *International Journal of Developmental Neuroscience, 20*(3–5), 459–465.
- Clarke, A., & Clarke, A. (2004). *Human resilience: A fifty year quest*. Philadelphia: Jessica Kingsley.
- Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7*(3), 249–253.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.
- Cole, K. N. (1995). Curriculum models and language facilitation in the preschool years. In M. E. Fey, J. Windsor, & S. F. Warren (Eds.), *Language intervention: Preschool through the elementary years*.
- Fey, M. E., Cleave, P. L., Ravida, A., Long, S. H., Dejmaj, A., & Easton, D. (1994). Effects of grammatical facilitation on the phonological performance of children with speech and language impairments. *Journal of Speech and Hearing Research, 37*, 594–607.
- Fey, M. E., & Proctor Williams, K. (2000). Recasting, elicited imitation and modeling in grammar intervention for children with specific language impairments. In D. V. M. Bishop (Ed.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 177–194). Philadelphia, PA: Psychology Press.
- Fudala, J., & Reynolds, W. (1986). *Arizona Articulation Proficiency Scale, 2nd ed.* Los Angeles: Western Psychological Services.
- Ingram, D. (1989). *Phonological disability in children*. London: Cole and Whurr.
- Kent, R. D., Weismer, G., Kent, J., & Rosenbek, J. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of Speech & Hearing Disorders, 54*, 482–499.
- Koegel, R. L., Camarata, S., Koegel, L. K., Ben Tall, A., & Smith, A. E. (1998). Increasing speech intelligibility in children with autism. *Journal of Autism and Developmental Disorders, 28*(3), 241–251.
- Leonard, L. B., McGregor, K. K., & Allen, G. D. (1992). Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research, 35*(5), 1076–1085.
- Long, S., & Fey, M. E. (1995). *Computerized profiling, Version 8.0*. Cleveland, OH: Case Western University.
- Miller, J., & Chapman, R. (1993). *SALT: Systematic analysis of language transcripts*. Madison, WI: University of Wisconsin.
- Nelson, K. (1989). Strategies for first language teaching. In M. Rice & R. L. Schiefelbusch (Eds.), *The teachability of language* (pp. 263–310). Baltimore: Brookes.
- Nelson, K. E., Carskaddon, G., & Bonvillian, J. D. (1973). Syntax acquisition: Impact of experimental variation in adult verbal interaction with the child. *Child Development, 44*(3), 497–504.
- Roid, G., & Miller, L. (1997). *Leiter International*

- Performance Scale—Revised*. Wood Dale, IL: Stoelting, Co.
- Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech & Hearing Disorders, 47*, 256–270.
- Shriberg, L. D., Kwiatkowski, J., Best, S., Hengst, J., & Terselic-Weber, B. (1986). Characteristics of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders, 51*(2), 140–161.
- Smith, A., & Camarata, S. (1999). Increasing language intelligibility of children with autism within regular classroom settings using teacher implemented instruction. *Journal of Positive Behavior Intervention, 1*, 141–151.
- St. Louis, K., & Ruscello, D. (1987). *Oral speech mechanism screening exam—Revised*. Austin, TX: Pro-Ed.
- Stone, W. L., & Yoder, P. J. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism, 5*(4), 341–361.
- Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tabachnick, B., & Fidell, L. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
- Tallal, P., Stark, R. E., & Mellits, D. (1985). The relationship between auditory temporal analysis and receptive language development: Evidence from studies of developmental language disorder. *Neuropsychologia, 23*(4), 527–534.
- Yoder, P., Camarata, S., & Gardner, E. (2005). *Predictors of later length of utterance and intelligibility in children with co-morbid speech and language impairment*. Manuscript in preparation. Vanderbilt University, Nashville, TN.
- Yoder, P., & Gardner, E. (2004). *Familiarity effects on estimates of MLU and intelligibility in preschoolers with phonological and grammatical impairments*. Submitted for publication.
- Yoder, P. J., Hooshyar, N., Klee, T., & Schaffer, M. (1996). Comparison of the types of child utterances mothers expand in children with and without Down syndrome. *American Journal of Intellectual Disability Research, 40*, 557–567.
- Yoder, P. J., Kaiser, A. P., & Alpert, C. L. (1991). An exploratory study of the interaction between language teaching methods and child characteristics. *Journal of Speech & Hearing Research, 34*(1), 155–167.
- Yoder, P. J., Kaiser, A. P., Goldstein, H., Alpert, C., Mousetis, L., Kaczmarek, L., & Fischer, R. (1995). An exploratory comparison of milieu teaching and responsive interaction in classroom applications. *Journal of Early Intervention, 19*(3), 218–242.
- Yoder, P. J., & McDuffie, A. (2002). Treatment of primary language disorders in early childhood: Evidence of efficacy. In P. Accardo (Ed.), *Disorders of language development* (pp. 151–177). Baltimore: York.
- Yoder, P. J., & Warren, S. F. (1998). Maternal responsiveness predicts the prelinguistic communication intervention that facilitates generalized intentional communication. *Journal of Speech, Language, and Hearing Research, 41*(5), 1207–1219.
- Yoder, P. J., & Warren, S. F. (2001a). Intentional communication elicits language-facilitating maternal responses in dyads with children who have developmental disabilities. *American Journal on Mental Retardation, 106*(4), 327–335.
- Yoder, P. J., & Warren, S. F. (2001b). Relative treatment effects of two prelinguistic communication interventions on language development in toddlers with developmental delays vary by maternal characteristics. *Journal of Speech, Language, and Hearing Research, 44*(1), 224–237.
- Zimmerman, I., Steiner, B., & Pond, R. (1992). *Preschool Language Scale-3rd Edition*. San Antonio: Psychological Corporation.

This research was supported by the Scottish Rite Foundation of Nashville, a P50 grant from NIDCD (P50DC03282) and the J. F. Kennedy Center Core grant from NICHD (HD15052). We are deeply grateful to the parents who trust us with their precious children. We also appreciate the hard and skillful work of Amiee Owen and Lisa Stepp, who conducted the treatment and transcribed the language samples. Finally, we appreciate the support from the larger P50 assessment core staff, which helped recruit participants and transcribe language samples and AAPS results (i.e., Mary Camarata, Stephanie Cofer, and Jennifer Croslin-Smith). Without all of this support, we could not do this research.

Correspondence concerning this article should be addressed to Paul Yoder, Peabody Box 328, Vanderbilt University, Nashville, TN 37203. E-mail: paul.yoder@vanderbilt.edu