JSLHR

Research Article

Treating Speech Comprehensibility in Students With Down Syndrome

Paul J. Yoder,[a] Stephen Camarata,[a] and Tiffany Woynaroski[a]

Purpose: This study examined whether a particular type of therapy (Broad Target Speech Recasts, BTSR) was superior to a contrast treatment in facilitating speech comprehensibility in conversations of students with Down syndrome who began treatment with initially high verbal imitation.
Method: We randomly assigned 51 5- to 12-year-old students to either BTSR or a contrast treatment. Therapy occurred in hour-long 1-to-1 sessions in students' schools twice per week for 6 months.
Results: For students who entered treatment just above the sample average in verbal-imitation skill, BTSR was superior to the contrast treatment in facilitating the growth of speech comprehensibility in conversational samples. The number of speech recasts mediated or explained the BTSR treatment effect on speech comprehensibility.
Conclusion: Speech comprehensibility is malleable in school-age students with Down syndrome. BTSR facilitates comprehensibility in students with just above the sample average level of verbal imitation prior to treatment. Speech recasts in BTSR are largely responsible for this effect.

Most people with Down syndrome (DS) experience lifelong difficulties in being understood by people outside of their immediate circle (Kumin, 2006). Even when compared with people with other intellectual disabilities matched by mental age, people with DS have far less comprehensible speech (Abbeduto & Murphy, 2004; Rosin, Swift, Bless, & Vetter, 1988). We will be referring to the disproportionate difficulty that students with DS have in producing speech that can be understood as a problem with speech comprehensibility.

Speech Comprehensibility as an Ecologically Valid and Functional Outcome of Speech Therapy for Students With DS

Speech comprehensibility is the extent to which an unfamiliar listener can understand what the speaker says (Hanson, Yorkston, & Beukelman, 2004). We were thoughtful and intentional in our selection of speech comprehensibility as the outcome of interest for speech therapy in school-age students with DS. We elected to use this terminology as opposed to the more commonly used term *speech intelligibility* because, at the measurement level, intelligibility has been defined differently depending on the research lab. For example, some researchers measure speech intelligibility exclusively in contexts in which the intended message is known because speech was elicited using a set of pictures, a list of words, or a set of sentences (for a review, see Kent, Miolo, & Bloedel, 1994). Compared with when they are labeling objects or pictures or answering simple questions, individuals with DS are less comprehensible when speaking in a conversational context in which the intended meaning is not known ahead of time by the listener (Farmer & Brayton, 1979). Thus, it is important that we judge the functional efficacy of speech therapy by assessing speech comprehensibility in conversations. In addition, we sought to study the extent to which an unfamiliar listener could fully comprehend the meaning of a student's spoken production.

For this reason, we opted not to focus on speech accuracy, which may be measured in conversations (e.g., as the percentage of consonants or phonemes correctly produced in speech samples) but focuses on the extent to which the student's production of words contains accurately produced speech sounds or phonemes relative to the adult target. For example, if a child produces the word "big" as [bI] (i.e., articulating the first consonant correctly, but omitting the second consonant entirely), the speech accuracy in terms of percentage of consonants correct would be 50% (Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997). Although one might assume that there is a high correlation

[a]Vanderbilt University, Nashville, TN
Correspondence to Paul Yoder: Paul.Yoder@vanderbilt.edu
Editor: Jody Kreiman
Associate Editor: Ewa Jacewicz
Received April 20, 2015
Revision received November 9, 2015
Accepted November 10, 2015
DOI: 10.1044/2015_JSLHR-S-15-0148

Disclosure: Paul J. Yoder and Stephen Camarata are the primary authors of the Broad Target Speech Recasts approach. Neither receives financial gain from the use of this treatment.

between speech accuracy and speech comprehensibility, the relationship is surprisingly weak.

In a sample of children with speech impairment for unknown reasons, only 16% of the variance in proportion of utterance attempts understood (a measure of speech comprehensibility) was accounted for by the percentage of phonemes correctly produced (a measure of speech accuracy), even though the two measures were derived from the same language samples (Shriberg & Kwiatkowski, 1982; Shriberg, Kwiatkowski, Best, Terselic-Weber, & Hengst, 1986). Thus, although related to some extent, speech accuracy is clearly distinct from speech comprehensibility. Speech comprehensibility is an ecologically valid and functional outcome of speech therapy for students with DS.

Speech comprehensibility is often conceptualized as a result of speech accuracy (Lagerberg, Åsberg, Hartelius, & Persson, 2014). That is, speech intervention often targets phoneme production with the implicit or explicit assumption that this, in turn, will improve speech comprehensibility (see Camarata, Yoder, & Camarata, 2006). However, this assumption is rarely examined in treatment studies of speech disorder, and given medium-sized correlations between measures of speech accuracy, speech intelligibility, and speech comprehensibility, a more direct focus on speech comprehensibility as an outcome is warranted, especially in populations, such as people with DS, with severe and persistent speech disorders.

Lack of Empirical Support for Speech-Therapy Effects on Speech Comprehensibility

There is surprisingly little high-level evidence that speech therapy improves speech comprehensibility in children with speech impairments (for a recent review, see Baker & McLeod, 2011). In previous studies, there has been very little internally valid evidence that speech therapy could specifically improve speech comprehensibility in children with DS. One problem is that we rarely directly test whether speech therapy actually enhances the comprehensibility of speech in children with DS, despite the fact that they are known to experience disproportionate deficits in this domain and the nearly universal agreement that the overarching goal of speech therapy is to improve children's ability to produce speech that can be understood by others in everyday conversations.

Broad Target Speech Recasts as a Potentially Useful Way to Enhance Speech Comprehensibility in Students With DS

An exception to the pattern of not directly testing whether speech therapy affects speech comprehensibility is the work on Broad Target Speech Recasts (BTSR; Camarata et al., 2006; Yoder, Camarata, & Gardner, 2005). A *speech recast* is an adult utterance that immediately follows a child's "platform" 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 but uses adult pronunciation. In this way, speech recasts provide indirect information about the accuracy of the child's production and provide models of accurate pronunciation, pitch, stress, and intonation. For example, if the child says, "Ah wa du," the speech recast might be "You want juice." The "broad target" part of BTSR is that the therapist uses speech recasts for any word the child attempts to say that is inaccurately produced and affords a developmentally appropriate recast.

BTSR might be particularly useful for students with DS because processable input for the adult pronunciation of words the child is trying to say is provided in a conversational context. The temporal proximity and semantic overlap of speech recasts with what the child is trying to say may make it easier for the child to compare the original utterance to the adult's recast. This contrast may make it easier to notice the phoneme, pitch, stress, rate, and intonation differences between the child's and the adult's utterances. Many instances of the heightened awareness of these differences increase the probability that the child will develop more accurate representations (i.e., memories) of the way words are said. We call these representations *sound templates*.

We use the term *sound template* instead of *phonological representation* because the speech recast provides more than just phonological information (Selkirk, 1981). Improvements to students' sound templates can occur even if students are not yet capable of producing the modeled speech sounds. The sound templates are thought to act as standards against which children judge and adjust their own attempts to produce the word. Accurate sound templates act as better guides for a child's own productions than inaccurate sound templates. Integration of the new information in the improved sound templates with existing semantic knowledge (as evidenced by the child's producing, or at least attempting, words that are recast) in contexts that resemble the child's most frequent communicative contexts (i.e., conversations) may increase the probability that newly learned information will become accessible outside of therapy in naturally occurring communicative periods and settings.

Two studies have previously been conducted using BTSR to improve speech comprehensibility. One of these studies examined effects of BTSR in a group of 52 preschoolers (mean chronological age of 3.7 years) with severe grammatical and speech impairments (Yoder et al., 2005). The results of this randomized controlled trial showed a moderate to large effect size ($d = 0.6$) for BTSR (in comparison to the control group) on generalized speech comprehensibility 8 months after the end of the treatment phase for a subgroup of children who entered treatment with the lowest raw scores on an articulation test that elicited production by showing pictures and asking the child to name the object. In the other study, BTSR was implemented with children who had DS and an average chronological age of 5.1 years. The results of this multiple-baseline across-participants experimental design supported an inference that there was a functional (i.e., causal) relation between the BTSR treatment and increases in generalized speech comprehensibility in four of six participants with

DS (Camarata et al., 2006). These results suggest that BSTR has the potential to improve the speech comprehensibility of at least some students with DS.

Rationale for Hypothesizing That BTSR Is More Optimal in Initially High Verbal Imitators

It would be helpful, from a clinical standpoint, to know which students with DS are most likely to achieve more optimal speech-comprehensibility growth in BTSR. In theory, for students with DS to benefit from a speech recast, they must retain the adult's recast and their own platform utterance in memory long enough to compare them, which is thought to make salient the information the recast adds. Also, the student must have sufficient oral motor ability to reproduce, or at least approximate, the adult model in future conversations. Thus, we suspect that students must have short-term auditory memory and speech motor control above some unknown threshold in order to benefit from BTSR.

Evidence suggests that some students with DS may not have the short-term auditory memory and/or motor control necessary to benefit from BTSR. Verbal short-term memory on average is shorter in children with DS than in typically developing children and those with developmental disabilities other than DS (Jarrold, Baddeley, & Hewes, 2000), even after controlling for IQ, mental age, and chronological age (Jarrold & Baddeley, 2001). Many professionals have also hypothesized that students with DS have speech-comprehensibility deficits because they lack the speech motor planning or control to execute their intended speech productions (Kumin, 2006; Miller & Leddy, 1999). There is variability on these factors within the DS population, however.

Identifying a measure that would be useful for determining which children with DS have above-threshold verbal short-term memory and motor ability to benefit from BTSR posed a challenge in this study. Short-term auditory memory and speech motor control, unfortunately, are difficult to measure in students with DS, particularly in students with DS who are younger or have more severe impairment. After much consideration, we concluded that a measure of verbal-imitation ability might reflect short-term memory and speech motor control. Verbal imitation is likely to require sufficient short-term auditory memory to retain a model long enough to imitate it and sufficient motor planning and execution ability to reproduce the adult model, at least fairly immediately.

Thus, to the extent that students' pretreatment tendency to verbally imitate reflects both sufficient verbal short-term memory to retain their own utterance and an adult recast in memory long enough to compare them and the speech motor ability to execute fine-tuned sound templates, one would expect BTSR to enhance speech comprehensibility in high verbal imitators better than low verbal imitators. This is an example of a *moderated treatment effect*. It would be seen as a statistical interaction between treatment-group assignment and pretreatment verbal-imitation ability predicting growth in speech comprehensibility in our sample of students with DS (Kraemer, Wilson, Fairburn, & Agras, 2002). Further detail regarding how we tested this moderated effect is provided later under Results.

Speech Recasts as an Active Component in BTSR

There is also value in determining the active component in BTSR that is responsible for the effect. Theory indicates that the most probable active component in BTSR is speech recasts (Camarata, 2010; Yoder et al., 2005). If the reason students assigned to the BTSR group have greater speech comprehensibility than students assigned to a contrast group is because the BTSR group had much greater exposure to speech recasts, then the treatment effect on speech comprehensibility should be explained, at least in part, by cumulative exposure to speech recasts. This is an example of a mediated treatment effect. Further detail regarding how we tested this mediated effect is provided later under Results.

Testing the hypothesized mediated treatment effect requires that we measure cumulative exposure to speech recasts in the BTSR group and in a contrast group to which students are also randomly assigned. Knowing when and in what contexts one should measure cumulative exposure to speech recasts in the contrast group is much easier when a contrast treatment is used as part of the research design. In addition, providing an active contrast treatment enables a more conservative test of the efficacy of BTSR than a vaguely defined business-as-usual control group.

Easy Does It as a Contrast Treatment

Many related-services personnel treating speech disorders in students use an approach to treatment that targets the phonological system using cuing strategies, placement techniques, and elicited production of a limited number of speech targets or phonological processes, rather than speech recasts. To provide a coherent, replicable approach that incorporated these strategies, we selected a well-defined, commercially available approach called Easy Does It (EDI; Drake, 2002). Using a research design that compares BTSR with a comparison therapy that both involves the type of methods that are commonly being used in the schools and is commercially available to clinicians affords a rigorous, clinically relevant test of the hypothesis that BTSR might work relatively better for students with DS who are initially high verbal imitators.

Research Hypotheses

In this study, we tested three hypotheses: (1) we suspected that there might be a significantly faster average slope of growth in generalized speech comprehensibility for students with DS assigned to the BTSR group than in students assigned to the EDI group (i.e., a main effect of treatment group). However, we thought it was more likely that (2) BTSR would facilitate speech comprehensibility growth, relative to EDI, to a larger extent in students with high rather than low pretreatment verbal-imitation ability (i.e., a moderated treatment effect). (3) We predicted that greater growth in speech comprehensibility from BTSR versus EDI would result from greater exposure to speech recasts in the BTSR group (i.e., a mediated treatment effect), at least for students with high pretreatment verbal-imitation ability.

Method

Participants

Fifty-one students with DS participated in the study. Inclusion criteria required that students (a) be between 5 and 12 years old at the time of pretreatment testing; (b) have been diagnosed with DS by a physician as reported by the parent; (c) have under 75% utterances comprehensible in a 20-min speech sample or be below the 10th percentile on the Arizona Articulation Proficiency Scale, Third Revision (AAPS; Fudala, 2000); (d) produce a minimum of 20 different comprehensible, referential words in a 20-min speech sample; (e) come from a home where English is the primary language; and (f) attend a school within a 45-min one-way car trip of the research center that gave permission for the research team to provide pull-out therapy during the student's school day. Exclusion criteria included (a) uncontrolled seizures; (b) diagnosis of attention-deficit disorder, autism spectrum disorder, or apraxia; (c) failure of hearing screening for both ears or wearing hearing aids in both ears; and (d) severely disruptive behavior as reported by the parent.

Table 1 includes means and standard deviations for several student descriptor variables, including the pretest on speech comprehensibility. As indicated, chronological age and verbal knowledge were significantly higher in the BTSR group than in the EDI group. All other pretreatment variables in Table 1 were nonsignificantly different between groups. Three students who were later diagnosed with ASD or who were provided with bilateral hearing aids were dropped from the analyses. Otherwise, no attrition occurred. Figure 1 provides the flow of interested parents and students through the design.

Table 1. Means and standard deviations of selected pretreatment variables by group.

Variable	BTSR	EDI
Chronological age (years)	7.8 (2.5)	6.5 (1.6)*
Total IQ score[a]	85 (18)	76 (14)
Verbal Knowledge score[a]	10 (3)	8 (3)*
Receptive Vocabulary standard score[b]	48 (20)	51 (15)
Number of different words produced across samples[c]	68 (38)	66 (42)
Mean length of utterance (in morphemes) across samples[c]	1.7 (0.5)	1.5 (0.4)
Weighted number of exact imitations[d]	14 (5)	12 (5)
Proportion of utterances comprehensible across samples[c]	.49 (.12)	.47 (.17)
Articulation standard score[e]	61 (8)	64 (8)

Note. BTSR = Broad Target Speech Recasts; EDI = Easy Does It. [a]Stanford–Binet Intelligence Scales, Fifth Edition. [b]Peabody Picture Vocabulary Test–Fourth edition. [c]Average score over two 20-min speech samples with an unfamiliar research staff. [d]Dynamic assessment of verbal imitation. [e]Arizona Articulation Proficiency Scale, Third Revision.

*$p < .05$.

Research Design

Students were randomized to either the BTSR or the EDI group using a computer program. The assessors were unfamiliar with the students outside of the assessment context, and speech-sample orthographers were unaware of group assignments. Speech comprehensibility was measured at four time points: pretreatment (Time 1), 2 months into treatment (Time 2), 4 months into treatment (Time 3), and immediately following completion of the 6-month treatment phase (Time 4). To prevent treatment contamination, different speech-language pathologists (SLPs) implemented BTSR and EDI. To control for therapist effects, the therapist who was initially implementing BTSR switched halfway through the treatment period to begin implementing the EDI treatment, and vice versa. The use of two active treatments reduced the probability of compensatory action that could account for results (e.g., seeking nonproject speech therapy to compensate for being assigned to the nonpreferred group).

Procedure

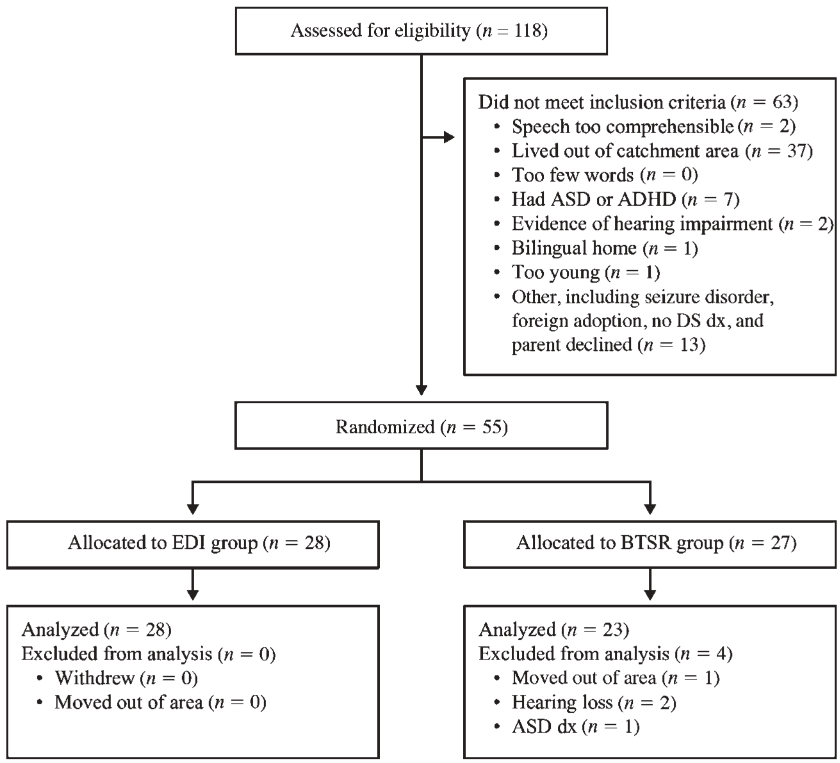
Overview

In the pretreatment period, a standardized measure of speech accuracy (the AAPS), a standardized measure of receptive vocabulary (the Peabody Picture Vocabulary Test–Fourth Edition; Dunn & Dunn, 2007), and a brief form of a standardized intelligence test (the Stanford–Binet Intelligence Scales, Fifth Edition abbreviated battery; Roid, 2003) were administered to characterize the students. Also during pretreatment, up to three administrations of the verbal-imitation measure and two 20-min speech samples were given. At Times 2–4, the two speech samples were repeated. At Time 4, the verbal-imitation measure was repeated. The setting for Time 1 assessments was the clinic. Table 2 summarizes the constructs, metrics, periods, and analytic roles of the variables used in the tests of the research questions.

Speech Samples (The Measurement Context for the Dependent Variable)

Two 20-min speech samples were obtained to measure speech comprehensibility and descriptive variables. The estimates of speech comprehensibility (and other descriptive variables) were averaged across two 20-min speech samples to increase the stability of the estimates compared with those from a single speech sample (Bruckner, Yoder, & McWilliam, 2006). The same sets of toys were used for all

Figure 1. Student recruitment, enrollment, randomization, and retention. ASD = autism spectrum disorder; ADHD = attention-deficit/hyperactivity disorder; DS = Down syndrome; dx = diagnosis; EDI = Easy Does It; BTSR = Broad Target Speech Recasts.



students, and these sets were not used during intervention sessions to afford tests of generalization across materials. Examiners were not the students' therapists, in order to afford for a test of generalization across communication partners. Examiners were responsive to students' actions and communication, asked questions about and delivered comments about the students' foci of attention and/or action, and refrained from imitation prompts and recasts. However, they did not ask topic-initiating questions or direct students to play in any particular way. In these ways, the speech samples were tests of generalization of speech-comprehensibility gains across several stimulus dimensions simultaneously. The sessions were video- and audio-recorded for later transcription.

Speech samples were transcribed and coded by research staff who were not the students' clinicians and who were

Table 2. Overview of variables by construct, procedure, period, role, and research question (RQ).

Construct	Procedure	Variable	Period	Role (RH)
Speech comprehensibility	Two 20-min conversational speech samples	Slope of the growth curve for the proportion of utterance attempts in which every word attempt was glossed, averaged across two 20-min speech samples	1–4	Outcome (1, 2)
Verbal imitation	Dynamic assessment of verbal imitation	Weighted sum of exactly imitated verbal stimuli	1, 4	Putative moderator (1)
Cumulative exposure to speech recasts	1. 25% of therapy sessions coded for use of speech recasts. 2. Record of duration of therapy sessions. 3. Record of student attendance to therapy sessions	Product of number of speech recasts per minute, average duration of therapy session, and total therapy sessions attended	Monthly during treatment phase	Putative mediator (2, 3)

equally unfamiliar with EDI and BTSR students. These staff were trained to a criterion of at least 80% agreement for three consecutive sessions and retrained when agreement fell below this criterion. Media files were controlled by ProcoderDV, a computer software program that records the time of occurrence of each child utterance and controls replay of utterances to aid transcription (Tapp & Yoder, 2003). Transcribers were allowed up to three passes per utterance. The speech samples were transcribed using Systematic Analysis of Language Transcripts (SALT; Miller & Chapman, 1990) conventions. The output was analyzed to derive lexical density (i.e., the number of word roots from all utterances), mean length of utterance (MLU) in morphemes from complete and fully intelligible utterances, and proportion of utterance attempts that were fully glossed from all utterances (i.e., speech comprehensibility). Transcription manuals are available from the first author.

Interobserver reliability was estimated on a random sample of at least 20% of all speech samples at all periods. Intraclass correlation coefficients were used as the estimate of reliability because they include information regarding differences between observers and variability among students in the reliability estimate (Yoder & Symons, 2010). The mean intraclass correlation coefficient across periods for speech comprehensibility was .86 (SD = .06). Observer drift was checked throughout each student's study period on a random selection of at least 20% of the speech samples. Utterance-by-utterance discrepancy discussions were held when agreement fell below 80%. The validity of the proportion of utterance attempts that were fully glossed is supported by its strong association (r > .80) with the average of four observers' real-time global rating of speech comprehensibility for same and different speech samples in 20% of the students at Time 4 (Yoder, Woynaroski, & Camarata, in press).

Dynamic Assessment of Verbal-Imitation Procedure (The Measurement Context for the Putative Moderator)

Student's pretreatment verbal-imitation ability was assessed using an adaptation of the dynamic assessment procedure of verbal-imitation ability (DAVI) from Yoder and Layton (1988). The goal of this assessment was to quantify the extent to which each student could be taught to verbally imitate 10 developmentally sequenced items (*aah*, *night-night*, *ma*, *pa*, *hi*, *bye-bye*, *cookie*, *she wants more*, *his mother opens the box*, *Susan is brushing her teeth*). Each item could be presented for up to three trials per session in up to three sessions (for a total of up to nine total trials per item). In each trial, the examiner presented the item (e.g., aah) and provided the student with an opportunity to approximate the target item (e.g., "You say it!"). If the student produced any vocalization after the command, verbal praise was provided (e.g., "Good talking!"). If the student approximated the model, an edible or tangible reward could be provided along with verbal praise (e.g., "You said it right!"). An approximation of an item was defined as a production that involved at least one correct consonant or typical consonant substitution plus at least one correct vowel produced in the correct position for each syllable or word of an item. An exact imitation was defined as a production that involved every consonant (i.e., no substitutions) plus every vowel produced in the correct position for each syllable or word of an item. If an approximation was given by the student on any trial, the item was retired. If an approximation was not given, the item was retained until the maximum of nine total trials was reached. *Verbal imitation* was the sum of exact imitations of the model weighted by how many trials the student was presented before exactly imitating the stimulus. Exact imitations were weighted 3, 2, and 1, respectively, after one through three, four through six, and seven through nine trials. Responses on this task were scored live.

Treatments

Therapy sessions were delivered twice a week for 6 months. The target duration for therapy sessions was 1 hr. The interventionists were project-trained, licensed, and certified SLPs, who were trained to criterion accuracy and monitored for adherence to the target treatment approach throughout the treatment phase. Treatment occurred at the students' schools.

EDI

This speech therapy method is based on the Cycles Approach proposed by Hodson and Paden (1983) and incorporates several methods commonly used in speech therapy. Treatment targets individually selected, deficient, but stimulable phonological patterns (e.g., syllableness, production of consonant singletons, voicing). Three phonological patterns are targeted per cycle (i.e., approximately 10- to 12-week period). At the end of each cycle, the student is reevaluated using tests of articulation (AAPS) and phonology (Hodson Assessment of Phonological Patterns–Third Edition; Hodson, 2004) to select targets for the subsequent cycle. As indicated by retesting, the new cycle may continue with the same pattern, if still deficient, or change focus to target new patterns.

In each EDI session, therapists conducted learning (1–3 min) and practice (the rest of the session) activities. Learning activities included the use of cuing and other techniques intended to establish accurate articulator placement and production of exemplars for the targeted process. In practice activities, therapists elicited production of words that incorporated the targeted phonological pattern during play with toy sets, in shared book reading, or with flashcards. This practice was designed to promote successful responses through requests for imitation of adult models and feedback on production accuracy (other than speech recasts). The flashcards, which include words for each level of each pattern, are provided by the EDI program. Word choices were made with careful consideration to their phonetic contexts. To standardize the method, we taught SLPs to use an average of one elicited imitation prompt per minute. To differentiate treatments, we excluded the auditory-bombardment listening activities that are typically included in the EDI treatment package and asked SLPs to refrain from using speech recasts.

BTSR

SLPs were taught to use at least four speech recasts per minute. However, a successful recast cannot be delivered unless (a) the student talks and (b) the student's intended message is understood. To enhance the probability that students would use a sufficient number of platform utterances that SLPs could sufficiently understand to provide accurate speech recasts, SLPs used verbal routines (Yoder & Davies, 1992; Yoder, Spruytenburg, Edwards, & Davies, 1995) and topic-continuing questions (Yoder, Davies, & Bishop, 1994; Yoder, Davies, Bishop, & Munson, 1994). Verbal routines and topic-continuing questions increase student talking. In addition, the predictability of the content in spoken turns in these routines provides sufficient listener support for interpreting the child's communicative message, despite poor speech accuracy. Unlike with the EDI approach, SLPs using the BTSR method were trained not to use elicited imitation prompts when delivering treatment.

Fidelity of Treatment Implementation (Measurement of the Putative Mediator)

Research staff collected records of (a) the duration of each session and (b) students' attendance to offered sessions. In addition, half of the treatment sessions were video-recorded, and therapists did not know which would be assessed for fidelity. Later, a random sample of 25% of the treatment sessions (i.e., 50% of the taped sessions) was selected for coding. A timed-event behavior-sampling method was used by trained observers to count the time of occurrence and number of times the therapist (a) provided a model and asked the student to say a word or syllable with a targeted sound (i.e., used an imitation prompt) and (b) used a speech recast. The *cumulative exposure to speech recasts*, the putative mediator, was the product of (a) the average rate of speech recasts per minute across the 12 coded sessions, (b) the average duration of therapy session in minutes, and (c) the total number of therapy sessions attended. Interobserver agreement on coded variables was checked on a random sample of 20% of the fidelity-of-treatment sessions that were coded by an independent coder. The primary coder did not know which session was going to be selected for reliability coding. The mean reliabilities for number of imitation prompts and number of speech recasts were .98 (SD = .03) and .95 (SD = .04), respectively.

Results

Preliminary Results

Univariate Normality of Primary Variables

Because the analysis methods used to test the research questions assume multivariate normality and because associations or differences on highly skewed (i.e., < |.8|) or kurtotic (< |3|) variables are more likely to violate this assumption, the distributions for all variables were examined. No analyzed variable, except for MLU at Time 1, violated these standards. Therefore, original scales were used for all variables but MLU. Log 10 transformation of the MLU variable brought its distribution within acceptable limits.

Pretreatment Between-Groups Equivalence

Even though random assignment increases the probability that groups will be nonsignificantly different on pretreatment variables that might explain subsequent between-groups differences in speech-comprehensibility growth, testing whether equivalence was achieved strengthens our confidence in any main or moderated treatment effects that we may observe. As shown in Table 1, BTSR and EDI treatment groups differed on two variables at Time 1 (age and verbal knowledge). However, these two variables had nonsignificant correlations associated with growth of speech comprehensibility, ps ≥ .5. In addition to the variables in Table 1, three other variables that were measured at Time 1 and thought to potentially be associated with growth in speech comprehensibility (i.e., raw score on the AAPS, proportion of single-word utterances that were comprehensible, and proportion of multiword utterances that were comprehensible) were nonsignificantly different between groups. Therefore, there was no evidence that pretreatment differences between groups could explain a main effect or a moderated effect of treatment on speech-comprehensibility growth.

Fidelity of Treatment and Differences Between Groups on Treatment Variables

Therapists met or exceeded their target rate of presumed active components in their respective treatments. The average rate of speech recasts in BTSR was 4.07 per minute (SD = 0.75), and the average rate of imitation prompts in EDI was 2.25 per minute (SD = 0.59). In addition, therapists were successful in not using the presumed active component of the model they were not assigned to use. The average rate of speech recasts in EDI was 0.01 per minute (SD = 0.02), and the average rate of imitation prompts in BTSR was 0.009 per minute (SD = 0.008). Thus, treatment contamination was negligible. Therefore, there was large differentiation between treatment models on the average rate of speech recasts, $d = -5.37$, and on the average rate of imitation prompts, $d = 7.65$.

The proportion of offered sessions that students attended did not differ by group—BTSR: $M(SD)$ = 68% (4%); EDI: $M(SD)$ = 71%(7%)—$t(49) = 1.04$, $p = .3$. However, the average duration of treatment sessions favored the BTSR group—BTSR: $M(SD)$ = 56 min (2.5); EDI: $M(SD)$ = 53 min (4.4)—$t(49) = -3.4$, $p = .001$. The average duration of session was, fortunately, unrelated to speech-comprehensibility growth, $t(50) = -0.67$, $p = .5$. Therefore, between-groups differences on average duration of session could not explain main or moderated treatment effects on growth in students' speech comprehensibility.

Modeling Growth in Speech Comprehensibility

Mixed-level modeling was conducted with the Hierarchical Linear Models software program (Raudenbush & Bryk, 2002). Mixed-level modeling began at Level 1 with the identification of the type of model that best fitted the data for students' growth in speech comprehensibility across the treatment period. At Level 1, speech-comprehensibility growth of each student was modeled over the four measurement periods (i.e., Times 1–4) with time periods nested within person, effectively creating individual growth curves for each student across the treatment phase.

Identifying the best-fitting growth curve involved comparing the most parsimonious model that enabled estimation of individual differences in the rate of growth to a more complex model that did so in a less straightforward manner. To unpack the process, the reader needs first to know that time in the study was centered at Time 1 so that the intercept of growth curves represented (and controlled for) students' speech comprehensibility at entry to the study. The simplest model that enables testing of the research questions is a random-intercept, random-slope model. This was compared to a random-intercept, random-slope, fixed quadratic model. A *quadratic* parameter quantifies acceleration and deceleration in growth (versus simply a fixed rate of growth) across the treatment period. A slope in models with quadratic terms is not the average rate of growth but instead the instantaneous growth rate at the intercept. *Fixed* parameters are the average of the individual growth curves for entry-level speech comprehensibility (intercept), slope, and/or acceleration in growth rate. *Random* parameters indicate the variability or individual differences on the aforementioned parameters across the treatment period. Formal tests for homogeneity of Level 1 variance and monitoring of covariance between factors were consistently conducted throughout model building to detect possible violations of statistical assumptions (i.e., heteroscedasticity or multicollinearity).

Best-Fitting Level 1 Model for Speech-Comprehensibility Growth

The random-intercept, random-slope model best fitted the data for speech-comprehensibility growth in students with DS across the study period (χ^2 for model fit = 356.8). This model substantively accounts for individual differences in both students' speech comprehensibility at entry to the study and students' growth rate in speech comprehensibility across the study period. More complex models that additionally considered acceleration or deceleration in student growth rate (i.e., that included a quadratic term) did not improve the fit to the data, $\chi^2(1) = 0.02$, $p > .5$. A significant fixed effect of slope in early modeling indicated that growth in speech comprehensibility occurred across the treatment period for students with DS on average, $t(49) = 5.9$, $p < .001$. A significant random effect of slope subsequently indicated that the rate of speech-comprehensibility growth significantly varied across students, $\chi^2(50) = 83$, $p < .001$. The latter effect means that there was variance in speech-comprehensibility growth across the treatment period at Level 1 that could be explained by students' assigned treatment group or the Treatment Group × Verbal Imitation interaction.

Primary Results

Absence of a Main Effect of Treatment Group on Speech Comprehensibility

After the best-fitting model of students' speech-comprehensibility growth was identified, we added the term for Treatment Group (reflecting BTSR vs. EDI group assignment) to evaluate whether the type of treatment that students received accounted for a significant amount of the variance in their speech-comprehensibility growth (i.e., to evaluate whether there was a main effect of BTSR treatment on average across all students with DS). The main effect of Treatment Group on the slope for speech-comprehensibility growth was nonsignificant and small, $t(49) = 0.88$, $p = .39$. Thus, growth in speech comprehensibility across the treatment period did not vary according to whether students received BTSR versus EDI treatment. Using pretest-to-posttest gain scores to aid comprehension of this noneffect, we find that the average gains were .11 (SD = .13) and .09 (SD = .11), respectively, in proportion of utterances fully glossed for BTSR and EDI groups, $d = 0.21$. Thus, on average, about .58 (SD = .15) of the students' utterances were comprehensible at Time 4. Table 3 contains the means and standard deviations for the observed speech-comprehensibility scores by group and period.

Presence of a Moderated Effect of Treatment Group by Verbal-Imitation Ability on Growth of Speech Comprehensibility

We added the Verbal Imitation predictor (reflecting students' verbal-imitation ability at Time 1) and the product term for Treatment Group × Verbal Imitation to determine whether the effect of Treatment Group on students' speech-comprehensibility growth varied according to how well students were able to imitate verbal models at entry to the study (i.e., to evaluate whether there was a moderated effect of BTSR treatment). The moderated effect of Treatment Group according to entry-level Verbal Imitation on the slope for speech-comprehensibility growth was significant, $t(47) = 3.27$, $p = .002$, pseudo-R^2 = .13. The coefficients and standard errors for the predictors in this final model are

Table 3. Means and standard deviations for speech comprehensibility for each period by treatment group.

Period	BTSR	EDI
1	.49 (.12)	.47 (.17)
2	.56 (.16)	.49 (.17)
3	.57 (.13)	.51 (.16)
4	.61 (.15)	.55 (.15)

Note. BTSR = Broad Target Speech Recasts; EDI = Easy Does It.

presented in Table 4. The significant Treatment Group × Verbal Imitation interaction on the slope means that students' speech-comprehensibility growth varied according to how well they were able to imitate verbal models at entry to the study. To be specific, speech-comprehensibility growth was greater with BTSR versus EDI for students with DS who were relatively higher in verbal imitation at Time 1.

Follow-up analyses indicated that students differentially benefitted from receiving BTSR versus EDI (i.e., experienced greater gains in speech comprehensibility in BTSR vs. EDI) if they entered treatment with a Verbal Imitation score of 15.5. This score is 0.5 of an *SD* above the sample mean of 13.1. The proportion of the sample scoring above this cut-point and thus being within the region of significance for this differential treatment effect was .53 (i.e., 27/51). It is important to note, however, that the findings for the moderated treatment effect are based on the entire group, not just the subgroup for which the differential treatment effect was observed.

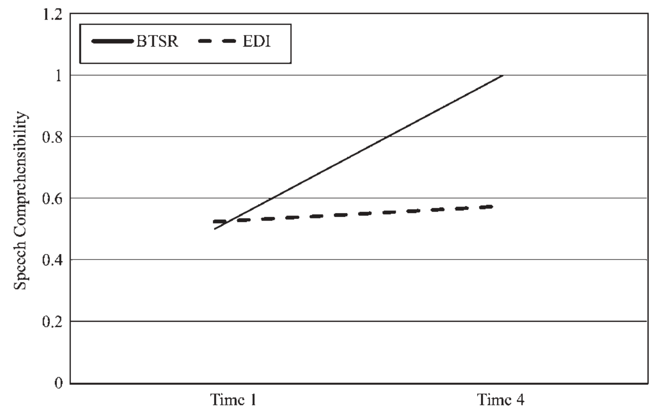
To illustrate this moderated treatment effect, we used predicted scores derived from the final model to generate growth curves for cases with scores of 15.5 for Time 1 Verbal Imitation assigned to receive BTSR versus EDI (see Figure 2). Growth curves represent the growth of the students who differentially benefited from BTSR versus EDI better than either observed scores at Time 4 or raw gain scores do (Singer & Willett, 2003). As shown in the figure, a student with a Verbal Imitation score of 15.5 would experience greater growth in speech comprehensibility in BTSR relative to EDI. Although both began the study with about .50 comprehensibility, the estimated Time 4 speech-comprehensibility outcome for a student with a score of 15.5 for Time 1 Verbal Imitation is almost 1.0 for BTSR and .59 for EDI. There was no interpretable Verbal Imitation value below which EDI was superior to BTSR.

Table 4. Unstandardized fixed and random coefficients and standard errors for predictors in the model testing the statistical interaction between Treatment Group and Verbal Imitation predicting growth of speech comprehensibility

Parameter	Coefficient (standard error)
Fixed effects	
For intercept	
Intercept	.4700*** (.0260)
Verbal Imitation	.0200** (.0060)
Treatment Group	.2900* (.1400)
Treatment Group × Verbal Imitation	−.0200* (.0090)
For slope	
Intercept	.0200*** (.0025)
Verbal Imitation	−.0016** (.0005)
Treatment Group	−.0270* (.0110)
Treatment Group × Verbal Imitation	.0024** (.0007)
Random effects	
Intercept	.0157*** (.0175)
Slope	.0001* (.0014)
Error	.0045 (.0093)

*$p < .05$. **$p < .01$. ***$p < .001$.

Figure 2. Growth curves for speech comprehensibility by group for students who scored 15.5 on the Time 1 Verbal Imitation scale. BTSR = Broad Target Speech Recasts; EDI = Easy Does It.



Cumulative Number of Speech Recasts Mediated the Group Effect in High Imitators

Because the Treatment Group effect was significant only for high imitators (i.e., those in the upper region of significance), we tested whether amount of speech recast was an active component in the BTSR treatment package only in the high-imitator subgroup. We tested whether the effect for BTSR treatment that we observed for high imitators was significantly reduced when we controlled for (i.e., accounted for) the cumulative number of speech recasts that students experienced over the course of the treatment period. In statistical terms, we tested whether the effect of Treatment Group on the slope of speech-comprehensibility growth was mediated through the cumulative number of speech recasts that students experienced.

A mediated effect is tested by determining whether the indirect effect of Treatment Group on growth in speech comprehensibility, through the cumulative number of speech recasts that students experienced, has a confidence interval with a lower bound that exceeds zero (see Figure 3; Hayes, 2009). This indirect effect, statistically speaking, is the product of the unstandardized coefficients for the effect of Treatment Group on cumulative number of speech recasts and the effect of cumulative number of speech recasts on slope of speech comprehensibility, controlling for Treatment Group.

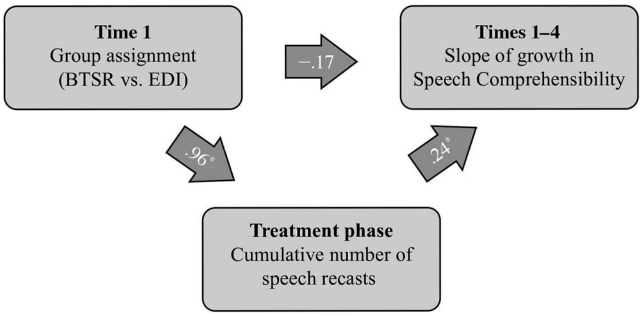
The cumulative number of speech recasts mediated the effect of Treatment Group on speech-comprehensibility growth, indirect effect = 0.035, 95% CI [0.0008, 0.0875]. The effect size of the indirect effect was large ($\kappa^2 = .35$).

Secondary Results

Time 1 Correlates of Verbal Imitation at Time 1

Of all measured Time 1 variables, six correlated with Time 1 Verbal Imitation, with *r* values ranging from .64 to .31. However, none of these correlates interacted with Treatment Group to predict slope of speech-comprehensibility

Figure 3. Treatment Group effect on slope of growth of speech comprehensibility was mediated through cumulative number of speech recasts, controlling for intercept (Time 1 speech comprehensibility). Coefficients are standardized. BTSR = Broad Target Speech Recasts; EDI = Easy Does It. *p < .05.



growth, all ps > .30. Therefore, none of the Time 1 variables that were correlated with pretreatment verbal-imitation ability could account for the Treatment Group × Verbal Imitation interaction on the slope of growth of speech comprehensibility.

Growth of Student Verbal Imitation Across Groups

To assess whether verbal imitation is a characteristic that changes in 6 months in students with DS, we tested whether significant growth occurred. It did, p = .001, d = 0.49. The means (and standard deviations) for Verbal Imitation at Times 1 and 4 across EDI and BTSR groups were 13.8 (4.65) and 15.4 (4.68), respectively. Therefore, the mean Verbal Imitation score at Time 4 was almost exactly the lower bound for the subgroup of students benefiting most from BTSR.

Discussion

This randomized controlled trial tested three a priori hypotheses. First, we predicted that if there were a main effect of treatment in this study, BTSR would facilitate speech comprehensibility more than a contrast treatment, EDI. Second, we hypothesized that BTSR, a treatment that involves speech recasts of child productions, should be more effective than EDI, a commercially available contrast treatment that relies in part on prompts for accurate imitation of adult models for exemplars of selected phonological processes, in students who already have relatively high verbal-imitation skill at the time they enter treatment. Third, we reasoned that the cumulative number of speech recasts that students experienced would be the mediator for the superior effect of BTSR in high verbal imitators. The latter two predictions were confirmed. Confirmation of a prior hypothesis is important because it suggests that the findings are more likely to be replicated than exploratory findings and it suggests a maturity in the cumulative knowledge base and theory that afforded the prediction (Lachenmeyer, 1970).

This epistemological discussion is relevant to why confirmatory conditional treatment effects (i.e., BTSR > EDI only in initially high imitators) are meaningful even when there is not a main effect of Treatment Group (i.e., BTSR is approximately equal to EDI without regard to pretreatment characteristics of students). The crux of the scientific and clinical issue is whether conditional treatment effects are likely to be replicated. A recent report indicates that confirmatory findings are more likely to be replicated than "surprising" results (Open Science Collaboration, 2015). We are thus proposing that the field needs to attend most to conditional treatment effects when they are predicted prior to study onset. With this in mind, it is useful to note that the conditional treatment effect seen in this study confirmed the prediction on all three of the primary dimensions of conditional treatment effects: the moderator, the end of the continuum of the moderator that would show the largest treatment effect, and the treatment group that would be superior in that region of significance.

The Relevance of the Findings for the Science Behind Speech Therapy

Any treatment effect, regardless of its conditional nature, on speech comprehensibility is noteworthy. Speech disorders in people with DS have proven to be both persistent and relatively resistant to treatment (Kumin, 2001, 2006). This challenge is not trivial, because the speech problems of children with DS are often quite severe and have an adverse impact not only on measures of articulation but also on the functional ability of these children to make the meaning of their spoken language understood (Kumin, 1994, 1996). The present study was designed to examine the effects of BTSR intervention relative to a contrast treatment on this latter aspect of this severe speech disability: functional comprehensibility in ecologically valid speech samples.

Superior gains in speech comprehensibility were seen in a subset of students who received BTSR. Some readers might find this result surprising because BTSR did not incorporate many of the strategies (e.g., cuing for accurate articulator placement and elicited imitation of adult models) that have long been considered staples or key components of intervention designed to improve speech disorders in DS (Kumin, 2006). Although Yoder et al. (2005) previously reported on improvements in speech comprehensibility for children with speech disorders without DS using BTSR, possible skepticism as to the viability of the approach for improving the severe speech disabilities seen in students with DS would be understandable. In fact, in our pilot work, BTSR was not clearly effective for all children with DS (Camarata et al., 2006).

In order to interpret these results, it is important to return to the distinction between speech accuracy and speech comprehensibility that was mentioned in the introduction. Much of the literature on assessment and treatment of speech disorders focuses on children with relatively high speech comprehensibility (Dodd, 2013).

The speech errors are often misarticulations (i.e., distortions) or patterns of consonant substitution that are noticeable in, and perhaps disruptive to, communication but not catastrophic with regard to comprehensibility (Dodd, 2013). Indeed, the literature is replete with reports on how to improve consonant production, even though it has long been known that vowel production provides the largest contribution to speech intelligibility and speech comprehensibility (Speake, Stackhouse, & Pascoe, 2012).

There is a high degree of face validity to the notion that teaching speech accuracy will readily translate to improvements in speech comprehensibility. After all, words consist of phonemes, and improving the accuracy of phoneme production should, logically, yield improvements in the ability of others to understand a child's intended message. Further, as also noted in the introduction, speech accuracy is often measured using consonant production with an inferential bridge to intelligibility (Lagerberg et al., 2014).

However, the following example illustrates why the correlation between speech accuracy and speech comprehensibility is only of medium effect size (i.e., $r = .4$) in children with severe speech disorders (Shriberg et al., 1986; Shriberg & Kwiatkowski, 1982). Consider the percentage of consonants correct in the production of the word "roll" as [wo] and as [o]. In both exemplars, there are no correct consonants, and the vowel is produced correctly. This yields a percentage of consonants correct of 0% in both cases. However, it stands to reason that listeners would be more likely to understand [wo] as an attempt to communicate the word "roll" than they would [o]. In addition, one can speculate that the gains in comprehensibility observed herein may also have related to relatively improved productions with regard to some phonetic features, even though remaining inaccurate at the phonemic level. In this example of "roll," perhaps BTSR facilitates a transition from a baseline form such as [o] to an improved, albeit still inaccurate, [wo]. This would likely correspond to improved speech comprehensibility, which would have been captured in our dependent variable. At the same time, this example underscores the fact that the current data cannot be interpreted as evidence that the students in this study who were high imitators improved their speech accuracy in BTSR better than EDI.

With speech comprehensibility as the dependent variable of interest in this study, we hypothesized that children who had sufficient verbal short-term memory and oral motor skills, as assessed by verbal imitation, to benefit from BTSR would show greater growth in BTSR than EDI for several reasons. First, BTSR provides exposure to more speech targets than EDI. Second, such recasts might enable the child to learn subphonemic phonetic features better in BTSR than EDI because the former does not hold the child accountable for phonemic accuracy, but EDI does. That is, asking the child to imitate may direct his or her attention to the phonemic level, rather than lower level phonetic information. Third, children may find it easier to process the linguistic information needed to improve sound templates from recasts than from imitation or elicited production prompts. Fourth, recasts provide the processable information in the context of minimally word-level units, which might enable integration of new linguistic information with other aspects of the language system better than phoneme-level imitation. Fifth, the similarity between BTSR therapy and conversation might enhance the probability of far transfer of newly learned knowledge from therapeutic interactions in the clinic to the unstructured conversations in the natural environment.

However, our hypothesis that BTSR would be effective for high verbal imitators was not a straw man by any means. In fact, we could also have constructed an argument that pretreatment faculty with imitation would have predicted better performance under EDI because children who are good imitators might be most responsive to a method that uses verbal imitation as a therapeutic method (Gillum, Camarata, Nelson, & Camarata, 2003). If both predictions were true, there would be no Treatment Group × Verbal Imitation interaction but rather only a Verbal Imitation main effect, predicting growth in speech comprehensibility. This did not occur.

Identifying the active components in treatment packages is important to helping us think clearly about how a treatment works. Finding that the effect of BTSR was mediated through cumulative exposure to speech recasts in high verbal imitators confirms that an active component in BTSR that drives gains in speech comprehensibility is cumulative number of speech recasts in this subgroup of students with DS. This mediated relation suggests that students with DS benefit from clinician models of correct word production (BTSR format) when they have sufficient verbal memory and motor capacity to compare sound templates and to alter their production when mismatches between child and clinician productions occur (Camarata, 2010; Dodd, 2013).

Clinical Ramifications

Speech comprehensibility can, and probably should, hereafter be at least one of our outcome measures of speech therapy, especially for children who are difficult to understand. The overwhelming majority of child speech-intervention studies focus on accurate phoneme production as an outcome measure, with an implicit or explicit presumption that this will translate to improved functional outcomes. Without criticizing speech accuracy as an important parameter, the results of this study suggest that clinicians should also attend to speech comprehensibility as an ecologically valid outcome measure, especially in children with severe speech disorders, such as those seen in DS.

The results of this study indicate that, for at least a subset of students with DS, speech accuracy training is not required to improve speech comprehensibility. This finding fundamentally challenges a foundational (and facially valid) assumption that has essentially dominated speech intervention for nearly a century (Kumin, 2006; Swift, 1918).

This is not to say that the results preclude the use of EDI or treatments that comprise similar strategies, but it is clear that some students with DS do not require treatment via conventional speech-therapy techniques to experience improvements on functional speech outcomes.

Although the current study did not test whether parents, classroom teachers, or paraprofessionals could be trained to implement BTSR effectively, we are optimistic about that prospect. One of the advantages of the BTSR treatment is that it does not require detailed professional knowledge of phonetics or articulation to implement. It simply requires that an adult use an interaction style that elicits frequent platform utterances from a student and have the mindfulness to provide regular recasts for a student's understandable but incorrectly produced utterances. We suspect that both caregivers and educators could be taught to incorporate BTSR into their everyday interactions with students with DS to support improved student speech comprehensibility in conversations. Future research is being planned along these lines.

Future Directions

There are a number of future directions for this line of inquiry. It is interesting that average verbal imitation grows in 6 months in students with DS. Although we have no evidence that our treatments facilitated this growth, the finding prompts the question of whether some other treatment can be used to facilitate verbal imitation, thus enabling more students with DS to benefit from BTSR. As mentioned earlier, studies are warranted that (a) explore the relationship between gains in speech comprehensibility and change on other acoustic and/or physiologic methods and (b) compare BTSR to other commonly used approaches in speech therapy. In addition, more comprehensive assessment of student characteristics that may moderate effects of treatment, such as verbal memory, speech processing, and speech motor control, would likely be illuminating. Also, the moderated and mediated treatment effects suggest that a future intervention study could potentially be conducted within the framework of "response to intervention" (Brown-Chidsey & Steege, 2010) using a SMART treatment design (Lei, Nahum-Shani, Lynch, Oslin, & Murphy, 2012), wherein treatment responders and nonresponders to BTSR are tested and the intervention adapted in various ways (e.g., intensifying BTSR, adding articulation drill) for those children who do not make gains in BTSR during the early part of the treatment phase. Future research is needed to determine whether the present results for BTSR extend to children with DS who differ in potentially important ways (e.g., are developmentally younger or older, have more severe intellectual impairment) relative to the students who participated in the present study.

Limitations

There are three important limitations to the present study. First, although sampling and coding of speech comprehensibility is an ecologically valid approach to studying the effect of speech treatment in students with DS, this study did not use perceptual or transcription methods to explore how treatment may affect other aspects of speech that are problematic for students with DS (e.g., speech accuracy) or use any acoustic and/or physiologic methods to examine possible correlates of speech-comprehensibility growth (e.g., average acoustic dispersion around known vowel targets or kinematics; for a recent review on this topic, see Kent & Vorperian, 2013). These represent important directions for future research.

Second, although EDI is a commercially available intervention that incorporates many strategies commonly used by SLPs in treating children with speech sound disorders (e.g., cuing strategies, placement techniques, elicited production) in everyday clinical practice (Brumbaugh & Smit, 2013; McLeod & Baker, 2014), it does represent only one of many approaches which SLPs may select for targeting speech comprehensibility. For example, it does not include as much direct training on articulator placement or as much "drill and practice" as some other commonly used approaches to speech treatment that are more firmly rooted in the theory that motor practice improves speech accuracy (Van Riper, 1947; Van Riper & Irwin, 1958). Likewise, EDI is not as solely focused on phonology as are some other methods, such as meaningful-minimal-contrast therapy (Blache, Parsons, & Humphreys, 1981; Weiner, 1981). In fact, we intentionally excluded one phonologically based aspect of the typical EDI treatment package (auditory-bombardment listening activities) in order to ensure a clear contrast with BTSR. EDI was used as an active control, which provides a more conservative test of the efficacy of BTSR than a no-treatment or business-as-usual control group. Thus, the results of the current study are best interpreted as evidence for the efficacy of BTSR in high verbal imitators, versus the superior efficacy of BTSR over the Cycles Approach or some other unified, theory-driven therapy method.

The last limitation of the study that we will review here was the use of verbal imitation as our proxy for verbal memory and motor ability. The developmentally young status of many of the students required the use of a proxy for these two abilities. The Verbal Imitation measure can be defended as a reasonable proxy for verbal memory and motor ability. However, the study would be strengthened if more direct measures of verbal memory and speech-production capability could have been included (Lanfranchi, Baddeley, Gathercole, & Vianello, 2012).

Strengths

However, there were several strengths to the current study. Among these strengths are (a) random assignment of students to groups, (b) verification that randomization yielded groups that were nonsignificantly different on multiple variables that were or could have been correlated with speech comprehensibility, (c) use of blind transcribers and coders, (d) low attrition, (e) high treatment fidelity, (f) use of an active comparison group, (g) counterbalancing of therapists with treatment methods, and (h) restriction of analyses to testing a priori predictions. Confirmation of a priori predictions signals a high probability of replication (Popper, 1963). Thus, we can have a high degree of confidence in the interpretability and replicability of the current study's findings for participants who are similar to those we studied.

Conclusion

Within the stated limitations, the current study provides several important scientific and clinical contributions. However, much further work is needed. We have just begun to examine whether speech therapy can facilitate the extent to which students' conversational utterance attempts are comprehensible. Nevertheless, the study contributes to the evidence base for treating speech disorders in students with DS using a recast approach, particularly those students with relatively higher imitation skills at the onset of intervention.

Acknowledgments

This research was funded by Institute of Education Science Grant R324A100225 (Principal Investigators: P. Yoder & S. Camarata) and supported by the National Institute for Child Health and Human Development through Vanderbilt Kennedy Center Grant P30HD15052 (Principal Investigator: E. Dykens). The project described was supported by CTSA award No. KL2TR000446 from the National Center for Advancing Translational Sciences (Principal Investigator: G. Bernard). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Science, the National Institute for Child Health and Human Development, National Center for Advancing Translational Sciences, or the National Institutes of Health.

We are very grateful to our wonderful staff (Elizabeth Gardner, Catherine Bush, Jenny Elrod, Amanda Johnson, Tricia Paulley, Rebecca Frey, Megan Ochab, and Meghan Weber), and the families who trust us with their precious children.

References

Abbeduto, L., & Murphy, M. M. (Eds.). (2004). Language, social cognition, maladaptive behavior, and communication in Down syndrome and fragile X syndrome. In M. L. Rice & S. F. Warren (Eds.), *Developmental language disorders: From phenotypes to etiologies* (pp. 77–96). Mahwah, NJ: Erlbaum.

Baker, E., & McLeod, S. (2011). Evidence-based practice for children with speech sound disorders: Part 1 narrative review. *Language, Speech, and Hearing Services in Schools, 42,* 102–139.

Blache, S. E., Parsons, C. L., & Humphreys, J. M. (1981). A minimal-word-pair model for teaching the linguistic significance of distinctive feature properties. *Journal of Speech and Hearing Disorders, 46,* 291–296.

Brown-Chidsey, R., & Steege, M. W. (2010). *Response to intervention: Principles and strategies for effective practice* (2nd ed.). New York, NY: Guilford Press.

Bruckner, C. T., Yoder, P. J., & McWilliam, R. A. (2006). Generalizability and decision studies: An example using conversational language samples. *Journal of Early Intervention, 28,* 139–153.

Brumbaugh, K. M., & Smit, A. B. (2013). Treating children ages 3–6 who have speech sound disorder: A survey. *Language, Speech, and Hearing Services in Schools, 44,* 306–319.

Camarata, S. (2010). Naturalistic intervention for speech intelligibility and speech accuracy. In L. Williams, S. McCleod, & R. McCauley (Eds.), *Interventions for speech sound disorders in children* (pp. 381–406). Baltimore, MD: Brookes.

Camarata, S., Yoder, P., & Camarata, M. (2006). Simultaneous treatment of grammatical and speech-comprehensibility deficits in children with Down syndrome. *Down Syndrome Research and Practice, 11,* 9–17.

Dodd, B. (Ed.). (2013). *Differential diagnosis and treatment of children with speech disorder* (2nd ed.). Chichester, United Kingdom: Whurr.

Drake, M. (2002). *Easy Does It for articulation: A phonological approach*. Austin, TX: Pro-Ed.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition: Manual*. San Antonio, TX: Pearson.

Farmer, A., & Brayton, E. R. (1979). Speech characteristics of fluent and dysfluent Down's syndrome adults. *Folia Phoniatrica et Logopaedica, 31,* 284–290.

Fudala, J. B. (2000). *Arizona Articulation Proficiency Scale, Third Revision*. Torrance, CA: Western Psychological Services.

Gillum, H., Camarata, S., Nelson, K. E., & Camarata, M. N. (2003). A comparison of naturalistic and analog treatment effects in children with expressive language disorder and poor preintervention imitation skills. *Journal of Positive Behavior Interventions, 5,* 171–178.

Hanson, E., Yorkston, K., & Beukelman, D. (2004). Speech supplementation techniques for dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology, 12,* ix–xxix.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76,* 408–420. doi:10.1080/03637750903310360

Hodson, B. W. (2004). *Hodson Assessment of Phonological Patterns–Third Edition*. Austin, TX: Pro-Ed.

Hodson, B. W., & Paden, E. (1983). *Targeting intelligible speech: A phonological approach to remediation*. San Diego, CA: College-Hill Press.

Jarrold, C., & Baddeley, A. (2001). Short-term memory in Down syndrome: Applying the working memory model. *Down Syndrome Research and Practice, 7,* 17–23.

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (2000). Verbal short-term memory deficits in Down syndrome: A consequence of problems in rehearsal? *The Journal of Child Psychology and Psychiatry, 41,* 233–244.

Kent, R. D., Miolo, G., & Bloedel, S. (1994). The intelligibility of children's speech: A review of evaluation procedures. *American Journal of Speech-Language Pathology, 3*(2), 81–95.

Kent, R. D., & Vorperian, H. K. (2013). Speech impairment in Down syndrome: A review. *Journal of Speech, Language, and Hearing Research, 56,* 178–210. doi:10.1044/1092-4388(2012/12-0148)

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59,* 877–883.

Kumin, L. (1994). Intelligibility of speech in children with Down syndrome in natural settings: Parents' perspective. *Perceptual and Motor Skills, 78,* 307–313.

Kumin, L. (1996). Speech and language skills in children with Down syndrome. *Mental Retardation and Developmental Disabilities Research Reviews, 2,* 109–115. doi:10.1002/(sici)1098-2779(1996)2:2<109::aid-mrdd9>3.0.co;2-o

Kumin, L. (2001). Speech intelligibility in individuals with Down syndrome: A framework for targeting specific factors for assessment and treatment. *Down Syndrome Quarterly, 6*(3), 1–8.
Kumin, L. (2006). Speech intelligibility and childhood verbal apraxia in children with Down syndrome. *Down Syndrome Research and Practice, 10,* 10–22.
Lachenmeyer, C. W. (1970). Experimentation: A misunderstood methodology in psychological and social-psychological research. *American Psychologist, 25,* 617–624.
Lagerberg, T. B., Åsberg, J., Hartelius, L., & Persson, C. (2014). Assessment of intelligibility using children's spontaneous speech: Methodological aspects. *International Journal of Language & Communication Disorders, 49,* 228–239.
Lanfranchi, S., Baddeley, A., Gathercole, S., & Vianello, R. (2012). Working memory in Down syndrome: Is there a dual task deficit? *Journal of Intellectual Disability Research, 56,* 157–166.
Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S. A. (2012). A "SMART" design for building individualized treatment sequences. *Annual Review of Clinical Psychology, 8,* 21–48.
McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics, 28,* 508–531.
Miller, J. F., & Chapman, R. S. (1990). *Systematic Analysis of Language Transcripts* (Version 1.3) [Computer software]. Madison: University of Wisconsin–Madison, Waisman Center, Language Analysis Laboratory.
Miller, J. F., & Leddy, M. (1999). Verbal fluency, speech intelligibility, and communicative effectiveness. In J. F. Miller, M. Leddy, & L. A. Leavitt (Eds.), *Improving the communication of people with Down syndrome* (pp. 81–91). Baltimore, MD: Brookes.
Open Science Collaboration. (2015, August 28). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.
Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge (essays and lectures)*. London, United Kingdom: Routledge & Kegan Paul.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.
Roid, G. H. (2003). *Stanford–Binet Intelligence Scales, Fifth Edition: Examiner's Manual*. Austin, TX: Pro-Ed.
Rosin, M. M., Swift, E., Bless, D., & Vetter, D. K. (1988). Communication profiles of adolescents with Down syndrome. *Communication Disorders Quarterly, 12,* 49–64.
Selkirk, E. O. (1981). On the nature of phonological representation. *Advances in Psychology, 7,* 379–388.
Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research, 40,* 708–722.
Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47,* 256–270.
Shriberg, L. D., Kwiatkowski, J., Best, S., Terselic-Weber, B., & Hengst, J. (1986). Characteristics of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders, 51,* 140–161.
Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrences*. New York, NY: Oxford University Press.
Speake, J., Stackhouse, J., & Pascoe, M. (2012). Vowel targeted intervention for children with persisting speech difficulties: Impact on intelligibility. *Child Language Teaching & Therapy, 28,* 277–295.
Swift, W. B. (1918). *Speech defects in school children and how to treat them*. Cambridge, MA: Riverside.
Tapp, J., & Yoder, P. (2003). Procoder (Version 1.1.2) [Computer software]. Nashville, TN: Vanderbilt University.
Van Riper, C. (1947). *Speech correction: Principles and methods*. New York, NY: Prentice-Hall.
Van Riper, C., & Irwin, J. V. (1958). *Voice and articulation* (2nd ed.). London, United Kingdom: Pitman Medical.
Weiner, F. F. (1981). Treatment of phonological disability using the method of meaningful minimal contrast: Two case studies. *Journal of Speech and Hearing Disorders, 46,* 97–103.
Yoder, P., Camarata, S., & Gardner, E. (2005). Treatment effects on speech intelligibility and length of utterance in children with specific language and intelligibility impairments. *Journal of Early Intervention, 28,* 34–49.
Yoder, P. J., & Davies, B. (1992). Do children with developmental delays use more frequent and diverse language in verbal routines? *American Journal of Mental Retardation, 97,* 197–208.
Yoder, P. J., Davies, B., & Bishop, K. (1994). Reciprocal sequential relations in conversations between parents and children with developmental delays. *Journal of Early Intervention, 18,* 362–379.
Yoder, P. J., Davies, B., Bishop, K., & Munson, L. (1994). Effect of adult continuing *wh*-questions on conversational participation in children with developmental disabilities. *Journal of Speech and Hearing Research, 37,* 193–204.
Yoder, P. J., & Layton, T. L. (1988). Speech following sign language training in autistic children with minimal verbal language. *Journal of Autism and Developmental Disorders, 18,* 217–229.
Yoder, P. J., Spruytenburg, H., Edwards, A., & Davies, B. (1995). Effect of verbal routine contexts and expansions on gains in the mean length of utterance in children with developmental delays. *Language, Speech, and Hearing Services in Schools, 26,* 21–32.
Yoder, P., & Symons, F. (2010). *Observational measurement of behavior*. New York, NY: Springer.
Yoder, P. J., Woynaroski, T., & Camarata, S. (in press). A psychometric study of a measure of speech comprehensibility in students with Down syndrome. *Journal of Speech, Language, and Hearing Research.*