

Research Article

The Efficacy of Recasts in Language Intervention: A Systematic Review and Meta-Analysis

Patricia L. Cleave,^a Stephanie D. Becker,^b Maura K. Curran,^c
Amanda J. Owen Van Horne,^c and Marc E. Fey^b

Purpose: This systematic review and meta-analysis critically evaluated the research evidence on the effectiveness of conversational recasts in grammatical development for children with language impairments.

Method: Two different but complementary reviews were conducted and then integrated. Systematic searches of the literature resulted in 35 articles for the systematic review. Studies that employed a wide variety of study designs were involved, but all examined interventions where recasts were the key component. The meta-analysis only included studies that allowed the calculation of effect sizes, but it did include package interventions in which recasts were a major part. Fourteen studies were included, 7 of which were also

in the systematic review. Studies were grouped according to research phase and were rated for quality.

Results: Study quality and thus strength of evidence varied substantially. Nevertheless, across all phases, the vast majority of studies provided support for the use of recasts. Meta-analyses found average effect sizes of .96 for proximal measures and .76 for distal measures, reflecting a positive benefit of about 0.75 to 1.00 standard deviation.

Conclusion: The available evidence is limited, but it is supportive of the use of recasts in grammatical intervention. Critical features of recasts in grammatical interventions are discussed.

Evidence-based practice, or evidence-informed practice, has been a central issue for the field of speech-language pathology for a number of years (Dollaghan, 2007). A key component of evidence-based practice is external research evidence, the most valuable of which comes from multiple studies that have been integrated in systematic reviews or meta-analyses. A number of systematic reviews relevant to language intervention with children have been conducted. Some reviews have evaluated the impact of language intervention in general (e.g., Cable & Domsch, 2011; Cirrin & Gillam, 2008; Law, Garrett, & Nye, 2004), including parent-implemented interventions (Roberts & Kaiser, 2011). These reviews have included a variety of interventions targeting various domains of language. Other reviews have had a narrow focus, evaluating multicomponent

interventions for areas such as school-age grammar (Ebbels, 2014), narrative (Petersen, 2011), or auditory processing skills (Fey et al., 2011). Although these types of reviews provide valuable information, because they combine different interventions or include multicomponent interventions, they do not allow for a determination of the efficacy of a particular intervention technique. The purpose of the systematic review and meta-analysis reported here was to identify and critically evaluate the research evidence available regarding the efficacy of a particular language intervention technique, conversational recasts, on grammatical development.

A *conversational recast* is a response to a child's utterance in which the adult repeats some or all of the child's words and adds new information while maintaining the basic meaning expressed by the child (Fey, Krulik, Loeb, & Proctor-Williams, 1999; Nelson, 1989). The additional information may be syntactic, semantic, and/or phonological. Under some systems, the term *recast* specifically refers to contexts in which the adult's response changes the voice or modality of the child's utterance. For example, a statement (Child: *Him need juice*) may be recast as a question (Adult: *Does he need some juice?*). In contrast, *expansion* refers to

^aDalhousie University, Halifax, Nova Scotia, Canada

^bUniversity of Kansas Medical Center, Kansas City

^cDeLTA Center, University of Iowa, Iowa City

Correspondence to Patricia L. Cleave: pcleave@dal.ca

Editor: Krista Wilkinson

Associate Editor: Carol Miller

Received July 22, 2014

Revision received December 8, 2014

Accepted January 9, 2015

DOI: 10.1044/2015_AJSLP-14-0105

Disclosure: The authors have declared that no competing interests existed at the time of publication.

episodes in which the adult's utterance maintains the child's words and basic meaning but modifies the child's sentence by changing structural or semantic details without changing the sentence modality (Child: *Him need juice*. Adult: *He needs juice*; cf. Fey, 1986). In the current study, following the lead of two key researchers of recasts, Nelson and Camarata, we included both types of recasts in a single category. *Recasts* are distinguished from *semantic extensions*, or *expatiations*, which are adult responses that continue the child's topic and add new information but do not necessarily contain any of the child's words (Child: *Dog running fast*. Adult: *He's in a hurry*). Recasts can be corrective, fixing an error in the child's utterance, or noncorrective, where the information added or modified is optional. Recasts may be simple, in which a single clausal element is added or modified, or complex, in which more than one clausal element is added or modified. In addition, recasts may be focused or broad. In interventions employing focused recasts, the recasts expand or correct the child's utterance in a way that provides the child with an example of one of her specific language goals. In interventions employing broad recasts, the recasts expand or correct the child's utterance along any dimension; there are no specific preidentified goals that must be included in broad recasts. Finally, the child is not required or even prompted to imitate the adult with either focused or broad recasting.

As discussed by Nelson and Camarata (Camarata & Nelson, 2006; Nelson, 1989), recasts are hypothesized to support language acquisition because they present feedback to the child in a way that highlights formal elements the child has not mastered. It is presumed that since the adult's recast immediately follows the child's utterance, the child notes the difference between her production and the adult's, which leads to acquisition of the language form. This happens most readily when the correction or addition by the adult represents a developmentally appropriate change. The temporal proximity and shared focus with the child's initial utterance increase the likelihood that the child will attend to the adult utterance. These features also reduce processing demands, allowing the child to make the critical comparison between her production and the adult's. In addition, the lack of a demand to produce the new structure is thought to free up cognitive resources for the comparison. Finally, because recasts are presented within an interactive context, the child's attention to the adult's language and motivation to communicate are enhanced (Camarata & Nelson, 2006; Nelson, 1989). Farrar (1992) provided evidence of the facilitative effects of corrective recasts. In that study, 2-year-old typically developing children were more likely to imitate a maternal use of a grammatical morpheme following a corrective recast than when the mother used the morpheme in another context. Saxton (1997) has argued that only corrective recasts influence development significantly, positing that corrective recasts not only present the child's target forms at opportune times but also provide negative feedback that the child's way of expressing meaning is not correct and show the child how to make the correction. In contrast, Leonard (2011) argues that recasts serve as a form of priming. The

adult's utterance introduces and reinforces the target form and need not be corrective in nature to be successful. Rather, it is the increased frequency of meaningful language input that is changing the child's language.

Language intervention programs have been studied in which recasts are the only or main technique used (Camarata & Nelson, 2006). These include ones in which broad recasts are used and ones in which focused recasts targeting the child's language goals are used. Recasts are also used in interventions as one of a set of techniques. These include interventions such as enhanced milieu teaching (e.g., Hancock & Kaiser, 2006) and the clinician-directed program of Fey and colleagues (Fey, Cleave, & Long, 1997; Fey, Cleave, Long, & Hughes, 1993).

Recasts are included intentionally or unintentionally in most forms of language intervention, but there is limited research on their efficacy. Therefore, it is not surprising that two research groups had independently planned literature reviews of recast studies using different but complementary methods. The first group (the first, second, and fifth authors) had embarked on a broad systematic review of studies of the impact of recasts on grammatical development, in which studies representing a wide range of experimental and nonexperimental designs were included. Only those studies were included in which recasts were identified by the authors as the sole or key active intervention component. The only additional treatment procedure allowed was what we refer to as *models*. Models are adult utterances that display the child's target but do not have the same contingent relationship to the child's utterance as recasts do. On the other hand, studies were excluded if the approach included prompts for imitation of adult recasts, as it is not possible to determine the unique contribution of recasts in these cases.

The second group (the third and fourth authors) was conducting a more quantitative meta-analysis of recast intervention studies. The meta-analysis included only experimental and quasi-experimental group studies that provided the necessary information to calculate comparable effect size. It allowed investigations that incorporated imitation or other techniques in recast-based intervention packages so long as they only treated children with specific language impairment. This decision allowed identification of a sufficient number of studies for analysis while making the comparison of the effect sizes meaningful. Because our initial approaches were different but complementary, we completed our studies as a team. Some differences in methods were maintained (e.g., search details, inclusion or exclusion of studies of children with intellectual disabilities) to protect the integrity of the method for research synthesis.

The experimental questions were as follows:

1. Are recasts more efficacious than either comparison interventions or no intervention in facilitating grammatical development among children with language impairment (LI)?
2. Are there features of the intervention that are associated with greater or more consistent effects?

Method

Searches

Systematic Review

To identify potential articles, we searched from 1970 to present using the databases and search terms in Table 1. Hand searches of the reference sections of all articles that passed all screenings were also conducted. New articles were screened using the same processes. The initial search was completed in February 2010, and it was updated in June 2013.

Two independent reviewers screened all articles first by title and then by abstract for relevance. For articles that passed, full papers were screened against the inclusion and exclusion criteria. After calculation of reliability, disagreements were resolved by consensus at each stage.

Primary criteria. Several criteria had to be met for a study to be included in our review and analysis: First, the article was published in a peer-reviewed journal between 1970 and 2013. Second, participants were children between the ages of 18 months and 10 years who spoke English. Studies involving children with LI or with intellectual disability (ID) were included. Those involving typically developing children were included only if the studies employed an experimental design. Third, the study identified recasts as a primary focus or primary component of the intervention. For recasts to be considered a primary component of the intervention, the authors must have identified recasts or expansions as a necessary part of the procedures or discussed the relationship between extent of expansion or recasts used and degree of progress associated with the intervention. Fourth, the study evaluated the effects of recasts at an early point on one or more measures of child grammar at a later point. This may or may not have included a control condition. Fifth, a pre-post measure of grammatical skills was employed, including broad, nonspecific indices of grammatical

skill such as mean length of utterance (MLU) and number of multiword utterances.

Additional criteria. Several additional criteria led to exclusion of potentially eligible studies: First, the participants had a hearing impairment, a diagnosis of autism spectrum disorder, or primary motor speech impairment. Second, recasts were being studied to facilitate acquisition of a second or later language. Third, the intervention included adult prompting for the child's imitation in addition to recasting.

Meta-Analysis

Table 1 includes the databases and search terms for the computer searches conducted in November 2013 for the meta-analysis. We screened all articles first on the basis of their titles and then, for those that passed the screening, against the inclusion and exclusion criteria, using the full articles. We also performed hand searches of the reference sections of included articles and of two systematic reviews related to language delay and treatment of language impairment (Ebbels, 2014; Law et al., 2004). The reference section of a third article (Zubrick, Taylor, Rice, & Slegers, 2007) was also searched, as it was known to include a significant number of relevant articles.

Articles were identified using the same inclusionary and exclusionary criteria as those used in the systematic review, with the following exceptions:

Inclusionary Criteria

- Studies of intervention packages involving prompts for imitation or expatiations were included if recasts were identified as a key component of the intervention.
- An experimental design that utilized recasts was compared using between-groups methods to a control

Table 1. Search results: Number of articles and percent agreement between raters.

Search	Total results	Included on basis of title	Included on basis of abstract and retrieved for examination	Retained
Systematic review: 2010 initial database search	2,982 (98%)	94 (77%)	46 (93%)	25
Systematic review: 2010 initial hand search	868 (95%)	72 (81%)	31 (97%)	5
Systematic review: 2013 update database search	687 (98%)	19 (84%)	8 (63%)	3
Systematic review: 2013 update hand search	109 (99%)	6 (83%)	0	0
Meta-analysis: 2013 initial database plus hand search	1,218	69 (98%)	54 (98%)	17
Meta-analysis: 2014 additional terms database plus hand search	99	15 (97%)	4 (100%)	1

Note. For the systematic review, the search terms for the first search were "recast* AND sentence AND language"; for the second search, they were "expansion* AND sentence AND language". Highwire Press was searched using the terms "(recast* OR expansions) AND sentence AND (grammar* OR morphosynta* OR synta* OR morph*) AND (treatment OR therapy) AND language AND child*". For the systematic review, all screenings and evaluations were completed independently by two reviewers and agreement was calculated. Disagreements were resolved by consensus at each stage. The meta-analysis used the EBSCO, PsycINFO, PubMed, LLBA, and Highwire Press databases. For the first search, the search terms were "'expressive language'+intervention NOT autism"; for the second search, "recasts+language+impairment" and for the third search, "'language impairment'+expansions NOT autism". An additional two searches were conducted in May 2013 using the terms "'milieu thereapy'+language" and "'milieu training'+language". For the 2013 search for the meta-analysis, initial reliability was established for the first 200 results; the primary coder independently assessed the title and abstract for the remainder of results. For the 2014 additional terms search, all results were coded by two researchers. For both, all articles retrieved in full were included or excluded via consensus.

group of participants, or using within-participant methods to a control structure or alternate intervention.

- Participants were children diagnosed as having specific LI or identified as late talkers; typically developing children were excluded, as were children with significant ID.
- Sufficient information to calculate effect size was included in the article or provided by an author upon request.

Exclusionary Criteria

- No group of participants was entered into the calculations twice. In cases where a series of articles discussed the same participant group at different stages during the intervention, the single article providing the clearest control group and largest number of participants was selected, and all other articles from the series were excluded from the meta-analysis.

These changes in inclusionary and exclusionary criteria from the systematic review were intended to permit meta-analysis of results from interventions using recasts as they are typically implemented in clinical practice. Because recasts in practice are most often implemented as part of a comprehensive therapy package, the effect sizes for recasts as part of a package were included in this meta-analysis. Studies that utilized recast-only interventions are labeled as such in each figure and are identified for discussion in the relevant sections of the text. The exclusion of participants with ID was intended to allow for meaningful computation of average effect sizes across all studies. Children with Down syndrome (Camarata, Yoder, & Camarata, 2006), autism (Grela & McLaughlin, 2006; Scherer & Olswang, 1989), or other developmental disabilities typically have especially severe language impairments and have been shown to respond differently to some intervention approaches than children with more specific disabilities (Yoder, Woynaroski, Fey, & Warren, 2014). It seemed questionable to add these children into average effect sizes mostly on the basis of children with specific LI. Thus, such studies were included in the systematic review, where they are discussed in qualitative terms, but they were not included in the meta-analysis. Where possible, *d* values and 95% confidence intervals (95% CIs) were calculated for the studies in the systematic review and not the meta-analysis; this information is available in Supplemental Tables S1 and S2.

Studies that utilized recasts as a key component of the intervention and focused on LI were not included in our meta-analysis if they lacked a clear comparison group or morpheme that did not rely on recasts. For example, Yoder, Molfese, and Gardner (2011) compared broad target recasts (BTRs) to milieu language teaching (MLT). Although this comparison is highly relevant to our questions and could possibly be taken as a comparison of a broad versus a focused recast approach, both groups received recasts as part of their intervention program. Therefore, the effect size

for this study is difficult to compare with those for studies contrasting a recast treatment and either no treatment or a treatment that does not contain recasts. Other studies that were excluded due to lack of a nonrecast comparison group examined important questions such as dose frequency (Bellon-Harn, 2012; Smith-Lock, Leitao, Lambert, Prior, et al., 2013) and cost-benefit analysis of different service provision models (Baxendale & Hesketh, 2003). Unfortunately, as in the Yoder et al. (2011) study, the comparisons were between two participant groups, each of which received recasts.

Finally, many potentially eligible studies did not provide a control group. Within-subject designs were included only if untreated grammatical targets were reported for comparison purposes. That is, none of the included studies were uncontrolled with effect sizes on the basis of mean pre-post differences within a single group.

Phase Assignment

The articles identified from both searches were categorized following the five-phase model for language intervention research described by Fey and Finestack (2009) to enhance comparability and to illustrate the developmental path that interventions based on recasts have taken. The first phase, Pre-Trial Studies, includes studies in which the focus is not intervention, so they were not included in our analyses. The second phase, Feasibility Studies, includes the earliest clinical trials. These explore issues such as the feasibility of an intervention and appropriateness of outcome measures. Although the impact of the intervention may be measured, the lack of a control comparison or a small sample size makes any conclusions regarding cause and effect tentative. All experiments involving only typically developing children were classified as feasibility studies, as the results did not speak directly to the efficacy of recasts for children with LI.

The Early Efficacy Studies phase represents the first phase in which it is possible to determine whether a cause-effect relationship exists between an intervention and an outcome. These studies are primarily concerned with internal validity and often sacrifice generalizability to maintain tight experimental control. These may include group or single-subject designs, but there must be some level of experimental control. For group designs, this involves a comparison of some sort: a contrast group for between-subjects designs or a contrast condition for within-subject designs. For single-subject designs, this involves treatment and control goals or replication across participants with staggered baselines. For our review, studies with outcome measures that involved assessment of intervention targets only (i.e., proximal measures) were classified as Early Efficacy Studies because they did not demonstrate generalization beyond the intervention targets. The next phase is Later Efficacy Studies. These studies also address the question of cause and effect but do so under more generalizable conditions and with a large enough sample size to ensure sufficient power. Studies that included outcome measures that involved more distal or

omnibus outcome measures, such as global measures from language samples or standardized tests, were included in this classification because the outcomes represented functional gains beyond direct intervention targets. The last phase is Effectiveness Studies. These studies explore whether the effects that have been seen in efficacy studies are seen under more typical, less controlled conditions. Studies in which speech-language pathologists or paraprofessionals conducted the intervention with their usual caseload in their usual setting were classified as Effectiveness Studies, as were parent programs that were delivered in a manner consistent with clinical services. See Fey and Finestack (2009) for a more detailed discussion of the framework.

Study Quality Rating

The same criteria were used to evaluate study quality for eligible articles from both the systematic review and the meta-analysis. Two reviewers rated each article independently, and disagreements were resolved by consensus. The 10 criteria used to determine study quality follow in the subsections below.

Participants. Were the participants adequately described, including information on age, expressive and receptive language skills, and cognitive skills?

Groups similar at start. Did the authors demonstrate that the groups did not differ statistically at the start? If the groups did differ, was this controlled for statistically (e.g., analysis of covariance), or was it determined that pre-experimental differences were not correlated with the outcome measures?

Therapy description. Was the use of recasts, including the goals of therapy and whether the recasts were broad or focused, adequately described?

Intensity of recasts. Was the intensity of recasts adequately described (i.e., session length, number of sessions, recast rate or frequency)?

Treatment fidelity. Was there evidence that therapy was provided as intended?

Blinding. Were the transcribers or coders unaware of group or goal assignment at a minimum?

Random assignment. Were participants or goals randomly assigned to treatment conditions? Studies in which there was a limitation on the randomization (e.g., participants were matched and then randomly assigned) were not rated as randomly assigned.

Reliability. Was the reliability of measures reported and adequate (i.e., reliability coefficients or interrater reliability $\geq .80$)?

Statistical significance. Were the statistical results adequately reported?

Significant effect. Were effect sizes reported or calculable for each variable compared?

Effect Sizes

Effect sizes were calculated for studies involved in the meta-analyses. Two meta-analyses were completed: one for

Early Efficacy studies and another for Later Efficacy and Effectiveness studies. Later Efficacy and Effectiveness studies were grouped together in the calculations because both report distal outcome measures. The method of comparison used within each study was identified as between subjects or within subject because this influenced calculation of effect sizes. For studies with multiple outcomes, only one outcome was selected for inclusion in each meta-analysis of overall effect size. In these cases, the outcome selected was that most relevant to the experimental questions and most closely tied to the goal of the specific intervention study. For studies with multiple nonoverlapping participant groups treated with interventions utilizing recasts, such as participants treated for production of two different grammatical targets (e.g., Leonard, Camarata, Brown, & Camarata, 2004) or studies comparing outcomes associated with different providers (e.g., Fey et al., 1993), each relevant participant group was entered into the calculation one time.

Effect sizes for studies utilizing between-groups methods were calculated using Hedges's g , following the methods of Turner and Bernard (2006). Hedges's g is similar to Cohen's d but contains a correction for the use of small samples with unequal group size. As a result of this correction, g is typically more conservative than Cohen's d , yielding smaller effect sizes for the same data. In line with Schmidt and Hunter's (2015) recommendation that the similarity across measures be reflected in notation, Hedges's g will be identified as d^* throughout this article.

Effect sizes for studies utilizing within-subject methods were calculated following the methods from Schmidt and Hunter (2015). The primary difference when analyzing outcomes from within-subject studies is that the observations are not independent; this is corrected for by inclusion of the correlation between measures in the equations. The relevant correlation is the intraclass correlation between intervention targets and control targets. Unfortunately, this is rarely reported. It was possible to calculate r for one study (Nelson, Camarata, Welsh, Butkovsky, & Camarata, 1996) due to provision of results for each participant, with $r = .77$.

It is expected that effect sizes will be largest when r is small because this implies that results are entirely due to the intervention itself. Given the paucity of information, we chose to be consistent in our use of r . The final weighted averages were carried out using an arbitrary correlation of .3 for all studies (which is a moderate-sized correlation) between pre- and posttest measures. The selection of this correlation was meant to balance the possibilities of Type I (falsely rejecting) and Type II (falsely accepting) error. McCartney and Rosenthal (2000) argue that it is just as problematic to reject an efficacious therapy as it is to use an ineffective one, particularly if the treatment choices are rather limited. If a higher correlation were used, such as the one reported by Nelson et al. (1996), the effect sizes would tend to be smaller and the CIs larger, making it more likely that we would determine that recasts are ineffective. Underestimation of the correlation within subjects may cause underestimation of variance and wrongly increase the

effect size; however, the error in d due to this lack of precision is likely to be small. Greater error likely arises from the limited information on test–retest reliability for experimenter-developed outcome measures (Schmidt & Hunter, 2015).

Effect sizes are considered significant if the 95% CI for d does not cross 0. Given significant results, positive effect sizes indicate a positive effect of recasts, and negative effect sizes indicate a detrimental effect of recasts. For averaging, effect sizes were weighted by the number of participants, such that studies with a larger number of participants contributed more heavily to the final effect size obtained. Weighted averages were then computed for both the Early Efficacy and Later Efficacy/Effectiveness outcomes. See the Appendix for additional details regarding effect-size calculations. Effect sizes are generally interpreted following Cohen's (1977) recommendations for defining small (0.2), medium (0.5), and large (0.8). Effect sizes fundamentally tell us how many standard deviations of difference there are between the treatment and the control groups. They are best understood in the context of the quality of the study, the particular outcome measures, and the treatment used. For example, it is important that the effort involved in the intervention be considered when evaluating the gain achieved; a small gain might be of value if little effort was involved. In addition, effect sizes might be small in an otherwise beneficial treatment because of measurement error (e.g., test–retest reliability) or because the comparison is between two highly effective interventions rather than between a highly effective intervention and a no-treatment control (McCartney & Rosenthal, 2000; cf. Yoder et al., 2011).

Results and Discussion

Searches

Information about which studies were included in the systematic review, the meta-analysis, or both is included in Table 2.

Systematic Review

The results from the computer and hand searches for 2010 and 2013 are detailed in Table 1. There were 2,982 articles identified in the original computer search. After title review, 94 remained, 46 of which remained after the abstract review. These articles were retrieved for study-level evaluation. Twenty-five articles were determined to meet all criteria. Hand searches of the reference lists resulted in an additional five articles. The updated search in 2013 found a total of 687 articles, which was reduced to 19 and eight, respectively, following the title and abstract reviews. Following a full article review, three additional articles were added to the systematic review. Hand searches of their reference lists identified no additional articles. When the results of the updated search were added to those of the original search, 33 articles were determined to meet the criteria set. An additional article (Baxendale & Hesketh, 2003) was identified

later, bringing the total to 34. One article (Baker & Nelson, 1984) contained two relevant studies, resulting in a total of 35 studies. These 35 studies formed the data set for the systematic review.

Meta-Analysis

As seen in Table 1, the searches conducted for the meta-analysis resulted in 901 citations, of which 832 were discarded on the basis of title or abstract. The full text of 69 articles was examined, and 14 articles were selected for inclusion. An additional four studies were identified as potentially eligible for inclusion but were discarded upon further examination due to lack of reporting of equivalent outcome measures across conditions or lack of reporting of sufficient data to calculate effect size. It should be noted that there were additional package-intervention studies in which the intervention was not described in sufficient detail to determine whether recasts were a key component of the intervention. Seven of the 14 studies were also included in the systematic review, and the other seven were package-intervention studies that were identified for the meta-analysis only. Two of the 14 (Fey et al., 1993; Leonard et al., 2004) contained multiple nonoverlapping participant groups; for each of these studies, two effect sizes are reported, one for each participant group.

Study Phase Assignment and Study Quality

The quality ratings for all studies are presented, grouped by phase, in Table 2. Interrater reliability on study quality rating was 87% for the systematic review and 97% for the meta-analysis. A description of the studies and summaries of their outcomes can be found in Supplemental Tables S1 and S2. In cases where package treatments used in the meta-analysis included outcome measures not relevant to grammatical skill, that fact is noted, and readers are directed to the relevant publication for further information.

Systematic Review

The evidence about the effect of recasting alone comes primarily from Feasibility (15) and Early Efficacy (12) studies. There are markedly fewer studies at the Late Efficacy (4) and Effectiveness (4) phases. As can be seen in Table 2, the study quality ratings varied greatly, ranging from 1 to 9. In general, the higher the quality rating, the more confidence a reader should have in the study's findings. Although there were high-quality studies in each phase, in general, a greater proportion of higher quality studies was found at the Early Efficacy and Late Efficacy phases, with 7/12 (58%) and 3/4 (75%) achieving at least 70% of relevant criteria, respectively. Only 4/15 (27%) of Feasibility studies and 1/4 (25%) of Effectiveness studies achieved this level. This pattern was not surprising. Feasibility studies are not designed to answer efficacy questions and thus typically involve less experimental control. Effectiveness studies are conducted in typical clinical settings, and achieving high levels of experimental control is thus more difficult. An

Table 2. Stage assignments, quality ratings, and study inclusion.

Citation	Participants	Therapy description	Intensity of recasts	Fidelity	Blinding	Random assignment	Reliability	Groups similar at start	Statistical significance	Significant effect	Quality rating	Study inclusion	Reason for exclusion
<i>Feasibility</i>													
Baker & Nelson (1984), Study 1	N	Y	N	N	N	N	N	N/A	N	N	1/9	SR	TD
Baker & Nelson (1984), Study 2	N	Y	N	N	N	Y	N	N	N	N	2/10	SR	TD
Baxendale & Hesketh (2003)	Y	Y	N	N	N	N	Y	N	Y	Y	5/10	SR	CG
Girolametto et al. (1999)	Y	Y	Y	N	N	N/A	Y	N/A	Y	Y	6/8	SR	RP
Hassink & Leonard (2010)	Y	Y	Y	Y	Y	N/A	Y	N/A	Y	Y	8/8	SR	RP
Hovell et al. (1978)	N	Y	Y	Y	N	Y	Y	N/A	N	N	5/9	SR	TD
McLean & Vincent (1984)	Y	Y	N	N	N	N	N	N/A	Y	N	3/9	SR	CG
Nelson (1977)	N	Y	Y	Y	N	N	Y	N	Y	N	5/10	SR	TD
Nelson et al. (1973)	N	Y	Y	N	N	N	Y	N	Y	Y	5/10	SR	TD
Pawlowska et al. (2008)	Y	Y	Y	Y	N	N/A	Y	N/A	Y	Y	7/8	SR	RP
Proctor-Williams et al. (2001)	Y	Y	Y	Y	N	N/A	Y	Y	Y	Y	8/9	SR	CG
Saxton (1998)	N	Y	Y	N	N	Y	N	Y	Y	Y	6/10	SR	TD
Scherer & Olswang (1984)	N	Y	N	Y	N	N	N	N/A	N	N	2/9	SR	TD
Schwartz, Chapman, Prelock, et al. (1985)	N	Y	N	N	N	Y	N	N	Y	N	3/10	SR	TD
Weistuch & Brown (1987)	N	N	N	N	N	N	N	N	Y	N	1/10	SR	ES
<i>Early Efficacy</i>													
Bradshaw et al. (1998)	N	Y	Y	Y	N	Y	Y	N/A	N	N	5/9	SR	SS
Camarata & Nelson (1992)	Y	Y	N	Y	N	Y	Y	Y	Y	Y	8/10	SR	SS
Camarata et al. (1994)	Y	Y	N	N	N	Y	Y	Y	Y	Y	7/10	Both	
Fey & Loeb (2002)	Y	Y	Y	Y	N	Y	N	Y	Y	N	7/10	SR	ES
Gillum et al. (2003)	Y	Y	N	N	N	Y	Y	Y	N	N	5/10	SR	ES
Leonard et al. (2004)	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	9/10	Both	
Leonard et al. (2006)	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	9/10	SR	RP
Leonard et al. (2008)	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	9/10	SR	RP
Loeb & Armstrong (2001)	N	Y	Y	Y	N	Y	N	N	N	N	4/10	SR	SS
Nelson et al. (1996)	N	Y	N	N	N	Y	Y	Y	Y	Y	6/10	Both	
Proctor-Williams & Fey (2007)	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	9/10	Both	
Schwartz, Chapman, Terrell, et al. (1985)	Y	Y	N	N	N	N	Y	N	Y	N	4/10	Both	

(table continues)

Table 2 (Continued).

Citation	Participants	Therapy description	Intensity of recasts	Fidelity	Blinding	Random assignment	Reliability	Groups similar at start	Statistical significance	Significant effect	Quality rating	Study inclusion	Reason for exclusion
Smith-Lock, Leitaó, Lambert, & Nickels (2013)	Y	Y	N	N	Y	N	N	Y	Y	Y	6/10	Meta	PI
Later Efficacy													
Fey et al. (1993), clinician-directed	Y	Y	N	N	N	Y	Y	Y	Y	Y	7/10	Meta	PI
Gallagher & Chiat (2009)	Y	Y	N	N	N	Y	N	Y	Y	Y	6/10	Meta	PI
Girolametto et al. (1996)	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	9/10	Both	
Roberts & Kaiser (2012)	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	9/10	Meta	PI
Robertson & Ellis Weismer (1999)	Y	Y	N	Y	N	N	Y	Y	Y	Y	7/10	Meta	PI
Tyler et al. (2011)	Y	Y	N	Y	N	N	Y	Y	Y	Y	7/10	Meta	PI
Tyler et al. (2003)	Y	Y	N	N	N	N	Y	Y	Y	Y	6/10	Meta	PI
Yoder et al. (2005)	Y	Y	Y	Y	N	Y	N	Y	Y	Y	8/10	SR	ES
Yoder et al. (2011)	Y	Y	Y	Y	N	Y	Y	Y	Y	N	8/10	SR	CG
Yoder et al. (1995)	Y	Y	N	Y	N	N	Y	N/A	N	Y	5/9	SR	SS
Effectiveness													
Camarata et al. (2006)	Y	Y	N	N	N	N	Y	N/A	N	N	3/9	SR	SS
Fey et al. (1993), parent program	Y	Y	N	N	N	Y	Y	Y	Y	Y	7/10	Both	
Fey et al. (1997), parent program	Y	Y	N	N	N	N	Y	Y	Y	N	5/10	SR	RP
Weistuch et al. (1991)	N	N	N	N	N	N	Y	N	Y	N	2/10	SR	ID

Note. N = no. Y = yes. SR = systematic review. N/A = not applicable. TD = participants were typically developing. CG = appropriate control group was not available. RP = participants were repeated from another study. ES = insufficient data to calculate effect sizes. SS = single-subject design. Meta = meta-analysis. PI = package intervention. ID = participants included children with intellectual disability.

improvement over time was noted with regard to study quality: More recent studies had better quality ratings, on average. Of the studies published in 2000 or later, 10/14 (71%) had quality ratings of 70% or above. For earlier studies, the proportion was 5/21 (24%).

Adequate participant descriptions were generally provided for studies including children with LI (81%; 22/27); no study involving children with typical development provided all required participant information. An adequate description of the therapy was included in all but the two studies by Weistuch and colleagues (Weistuch & Brown, 1987; Weistuch, Lewis, & Sullivan, 1991). However, the general therapy description criteria did not require that the intensity of recasts be presented. As can be seen in Table 2, only 17/35 (49%) studies either reported dosage information or provided sufficient information to calculate it. Randomization was included in 16/31 (52%) of studies where it was relevant. Blinding was the criterion least consistently scored. Only five studies, four of which involved the same research program (Hassink & Leonard, 2010; Leonard et al., 2004; Leonard, Camarata, Pawlowska, Brown, & Camarata, 2006, 2008) reported that the transcribers or coders were appropriately unaware of group or goal assignments.

Meta-Analysis

The recast-only studies used in the meta-analysis are included in the descriptions already given and were primarily classed as Early Efficacy studies (5/7) with within-group designs (5/7). The package studies, which were exclusively examined in the meta-analysis, were primarily Later Efficacy studies (6/7) utilizing between-groups designs (7/7). The control groups were primarily no-treatment or wait-list group, with one group receiving business-as-usual treatments (Roberts & Kaiser, 2012). Due to the specific inclusionary and exclusionary criteria needed to allow for effect-size calculation of outcomes for children with LI, all of the package studies necessarily met the criteria of adequate participant descriptions and adequate description of the therapy. However, other quality ratings were less reliably reported. Intensity of recasts was reported for only 1/7 (14%) of the studies added (Roberts & Kaiser, 2012), and blinding was similarly reported for only 1/7 studies (Smith-Lock, Leitao, Lambert, & Nickels, 2013). Treatment fidelity was adequately reported for 3/7 (43%) studies. A quality rating score of 70% or greater was achieved by 5/7 (71%) studies.

Question 1: The Effects of Recasting Intervention

Systematic Review

Our first research question focused on whether recast interventions were more efficacious than either comparison treatments or no treatment in facilitating grammatical performance among children with LI. Findings from Feasibility studies could not speak directly to this question, but a number of them did provide support for including recasts in language intervention. This came through two sources. One type of support came from studies that examined the impact of recasting with children with typical language development.

All eight studies reported positive effects (Baker & Nelson, 1984; Hovell, Schumaker, & Sherman, 1978; McLean & Vincent, 1984; Nelson, 1977; Nelson, Carskaddon, & Bonvillian, 1973; Saxton, 1998; Scherer & Olswang, 1984; Schwartz, Chapman, Prelock, Terrell, & Rowan, 1985).

The second type of support came from studies that explored the correlation between input involving recasts and a child's language development at a later point in time. There were three such studies, and they reported different results. Girolametto, Weitzman, Wiigs, and Pearce (1999) reported medium correlations between mothers' use of recasts and children's language development. Baxendale and Hesketh (2003) reported that the children of parents who recast at a higher rate made greater gains than those of parents who recast at a lower rate, regardless of whether the parents had been trained in the parent-training program or in individual sessions. Both of these studies involved only children with LI. However, Proctor-Williams, Fey, and Loeb (2001) found a significant positive relationship between target-specific parental recasts and children's language development for typically developing children but not for children with LI.

Twenty studies yielded evidence that bears directly on our first research question. Sixty percent (12/20) of the studies were Early Efficacy studies. Two early studies found that recasting produced better results than interventions in which child utterances were not recast (Bradshaw, Hoffman, & Norris, 1998; Schwartz, Chapman, Terrell, Prelock, & Rowan, 1985). Loeb and Armstrong (2001) also found support for recasting compared to no treatment, but there was no evidence that recasting was superior to modeling.

Camarata, Nelson, and colleagues have conducted a series of studies exploring the effects of recasting. In all four studies, a within-subject design was used in which grammatical morpheme targets were randomly assigned to either a recasting or an imitation condition. Camarata and Nelson (1992) and Camarata, Nelson, and Camarata (1994) found that fewer presentations were required before the child's first spontaneous productions in the recasting condition. This was true for a subset of children with poor preintervention imitation skills (Gillum, Camarata, Nelson, & Camarata, 2003). The study by Nelson et al. (1996) included a group of typically developing children matched on language level. It also included targets that were absent from the child's language and ones for which the child displayed partial knowledge. Results showed that recasting was superior to imitation for both groups of children and for both types of targets.

Leonard and colleagues have produced three articles from a study of a recast intervention. The first reports on the results after 48 intervention sessions (Leonard et al., 2004), the second after 96 intervention sessions (Leonard et al., 2006), and the third 1 month after the completion of the intervention (Leonard et al., 2008). In this study, children received the recast intervention for either auxiliary *be* or third person singular *-s*. Outcomes for untreated control goals were also reported. In the 2008 study, a comparison group receiving general language stimulation that involved

broad recasts was also included. All children started treatment at floor on their target and control goals. At all points in time, the children displayed better performance on their treatment goals than their control goals. Furthermore, greater gains were seen in the groups receiving focused recasts than the one receiving general language stimulation.

In contrast, the results of two studies did not provide support for recasting. Fey and Loeb (2002) provided a very specific type of noncorrective recast, in that a child's comment was responded to with a recast that was a question with the auxiliary in initial position. The growth in auxiliary production was compared for children who received the recasts and those who did not. No effect of recasting was seen in either typically developing children or those with LI. Fey and Loeb proposed that this lack of effect was related to the type of recast provided (i.e., auxiliary-fronted questions) and to the children's developmental level being too low to utilize the information in such recasts. Proctor-Williams and Fey (2007) used recasting to teach novel irregular past-tense verbs. Because the verbs were novel, any gains could be wholly attributed to the treatment. The researchers were particularly interested in recast rate, so they compared performance on verbs recast at a conversational rate (0.2/min) and an intervention rate (0.5/min) to ones that were only modeled. There was a positive effect for recasting at the lower rate for the typically developing children, but for those with LI, there was no difference between the conditions. The typically developing children actually performed significantly less well at the higher rate than they did at the lower rate. This suggests that at rates that appear to be too frequent for typical children, thus limiting performance, children with LI may still not get enough recasts to affect their learning.

Four of the studies were Late Efficacy studies, three of them conducted by Yoder and colleagues. In 1995, Yoder, Spruytenburg, Edwards, and Davies reported on a single-subject-design study involving four children with mild ID. For the three children who began the study at an early developmental stage, Brown's Stage I, there was evidence of an effect of the intervention on their MLU. The fourth child was at Stage IV at the beginning, and one of the other children achieved Stage IV during the study. There was no evidence that recasting increased MLU for children who were at Brown's Stage IV.

Yoder, Camarata, and Gardner (2005) used a group design in which 52 children with speech and language impairments were randomly assigned to an experimental intervention involving both speech and grammar broad recasts or to usual services. In the experimental condition, the clinicians employed grammatical recasts following well-articulated child utterances and speech recasts following ones which were poorly articulated. The researchers found no significant effects of broad grammatical recasts on the children's sentence lengths.

The third study (Yoder et al., 2011) compared broad target recasts (BTRs) to milieu language treatment (MLT). MLT makes frequent use of recasts along with a hierarchy

of prompts, including direct imitation prompts to target specific language goals. Using growth curve modeling, the researchers showed that both interventions resulted in positive changes. There was no difference between the interventions when all children were included. In fact, we calculated a small, nonsignificant negative effect size for BTR compared to MLT ($d = -0.09$, 95% CI [-3.08, 2.90]). However, on the basis of a moderator analysis, Yoder et al. found that children whose initial MLU was less than 1.84 at study onset made greater gains if they received MLT than if they received broad target recasting. For children with initial MLUs greater than 1.84, no difference in response to interventions was found.

Girolametto, Pearce, and Weitzman (1996) reported on the results of a program in which parents were taught to use facilitative techniques, including recasts. Recasts initially focused on target words and moved on to word combinations when the children demonstrated enough progress on vocabulary. The children in the experimental group displayed greater growth in word combinations than did a delayed-treatment control group.

There were four investigations that best fit criteria for Effectiveness studies, although they are far from prototypical examples. Three of the studies involved a parent-training program, and the fourth involved a study in which the staff administered the recasting intervention in a clinical setting. The two articles by Fey et al. (1993, 1997) reported on outcomes of a single intervention at two different times, after 5 months of treatment and after 10 months of treatment. Only the parent program was included in the systematic review, because the clinician-directed program involved an imitation component, making it difficult to disentangle imitation and recasting as therapeutic techniques. Parents were trained to use recasts and focused stimulation techniques for specific grammatical targets. After 5 months of treatment, the children in the treatment group displayed better grammatical development than a delayed-treatment control group. Recasting by parents in the parent program was compared to recasting by parents of children in a clinician-directed program who did not receive instruction on recasting. The two groups of parents did not differ in their use of recasts before intervention, but after 5 months of treatment, the parents in the parent program recast at a significantly higher rate. The second article explored the amount of gain made with a second 5 months of treatment. The children's scores were significantly higher at the end of 10 months of treatment than they were at the end of 5 months. Furthermore, the children whose parents produced a higher rate of recasting (i.e., at least 1.27/min) made greater gains than those whose parents recast at a lower rate (i.e., less than 0.67/min). Further bolstering confidence that the treatment was causing the change, the group of children who did not receive a second phase of treatment did not show significant growth in the 5 months following their dismissal.

The study by Weistuch et al. (1991) also involved a parent program. In this study, the children in the experimental condition showed greater increases in MLU than

a control group. It is important to note, however, that the recast rate of parents in the experimental group was not different statistically from that of parents in the control group. This suggests that the gains made in children's MLU may have been due to factors other than recasts. However, the researchers do report a significant association between changes in the mothers' recast rate and their children's growth in MLU for the experimental group but not the control group.

The final Effectiveness study was conducted by Camarata et al. (2006). Six children with Down syndrome participated in a study targeting both intelligibility and grammatical skills. Although five participants showed growth in MLU, in only two cases was there a clear separation between baseline and treatment phases, thus providing evidence of a treatment effect.

Because our criteria for inclusion in the systematic review were quite broad, the design of the studies and thus the type and strength of the evidence varied greatly. The 34 studies were also heterogeneous in terms of intervention details and recast types. Nevertheless, across all phases, the vast majority of studies provided support for the use of recasts in interventions targeting grammatical development.

Meta-Analysis

Early Efficacy studies. The results from the meta-analysis of Early Efficacy studies are presented in Figure 1. These seven studies showed a great deal of variation in both the measures used to evaluate performance and the effect sizes obtained. Note that Leonard et al. (2004) contributed two outcomes to this analysis, for a total of eight measures in the final average here.

Half of the effect sizes were significantly different from zero (Camarata & Nelson, 1992; Camarata et al., 1994; Leonard et al., 2004). All four are within-participant studies utilizing recast-only interventions, with two data points coming from the same article. The two groups reported in Leonard et al. (2004) both demonstrated significantly better performance on treated morphemes than on untreated control morphemes. This effect was significantly larger for the participants treated for third person singular *-s* ($d = 2.08$, 95% CI [1.72, 2.43]) than for those treated for auxiliary production ($d = 1.07$, 95% CI [0.75, 1.39]), a finding that the researchers attribute to dose effects.

The sole package intervention to be classified as Early Efficacy (Smith-Lock, Leitao, Lambert, & Nickels, 2013) had a small nonsignificant effect size ($d^* = 0.41$, 95% CI [-2.52, 3.34]). This study employed recasts as part of a treatment program which also included explicit teaching and cloze tasks and utilized a variety of providers.

When an average d was computed for the eight outcomes, an average effect size of .96 and 95% CI [0.76, 1.17] were found. This reflects a positive benefit of approximately 0.75 to 1.00 SD on outcome measures closely aligned with the treatments. All of the Early Efficacy outcomes exhibited positive d values, though actual magnitudes varied from 0.23 to 2.08. It is notable that all of the studies except those from the Leonard et al. (2004) article

compared treatments using recasts to an alternative language treatment approach. Larger effect sizes might be observed if the comparison were to a no-treatment control. Statistical comparisons of average outcomes for studies with differing control groups and varying quality ratings would be ideal but are not possible at this time due to the small number of studies available.

Later Efficacy/Effectiveness studies. The seven Later Efficacy/Effectiveness studies included five package interventions and two recast-only interventions. In contrast to the previous analysis, only one of these studies (Roberts & Kaiser, 2012) compared the experimental group to an alternate-treatment control group. Note that Fey et al. (1993) contributed two outcomes to the averages discussed here because of the study design; one participant group received recast-only intervention implemented by parents, and the other received a package intervention implemented by a clinician.

A forest plot presenting outcomes for each study and the weighted average can be observed in Figure 2. Four of the eight outcome measures were found to exhibit a positive effect size significantly different from zero. These include three package interventions (Fey et al., 1993, clinician group; Roberts & Kaiser, 2012; Robertson & Ellis Weismer, 1999) and one recast-only intervention (Fey et al., 1993, parent group).

Examination of Figure 2 reveals that the results from Girolametto et al. (1996) have an unusually large CI. This is partially due to our selection of multiword combinations as an outcome measure; both multiword combinations and structural complexity on the MacArthur-Bates Communicative Development Inventories (Fenson et al., 1993) are reported. The d^* for the index is 0.95 (95% CI [-3.56, 5.48]), whereas the multiword combinations have a smaller d^* of 0.68 and a larger 95% CI [-17.09, 18.45]. We selected multiword combinations from the language sample because they provided the most empirically derived measure. The MacArthur-Bates index, a parent-report measure, has a narrow range of scores and thus is less variable.

When an average d^* was computed for the seven studies, an average effect size of 0.76 and a 95% CI [0.46, 1.06] were found. This reflects a positive benefit of about 0.5 to 1.0 SD . All of the Late Efficacy studies exhibited positive d^* values.

Although the CIs for Early and Late Efficacy studies do overlap, the Early Efficacy studies have larger average effect sizes. Recall that effect sizes are best interpreted in context. Thus, those studies reporting outcomes most closely aligned with treatment seemed to show larger gains. This is despite the fact that most Early Efficacy studies were comparing recasts to alternative language treatments, whereas most of the Late Efficacy studies were comparing recasts to no-treatment controls. This suggests that when outcomes are measured in a way that is tightly aligned with therapy goals, greater gains are observed. However, even when outcome measures that are only broadly related to the targets are used, gains are observed, suggesting some generalization across language structures.

Figure 1. Forest plot of Early Efficacy studies.

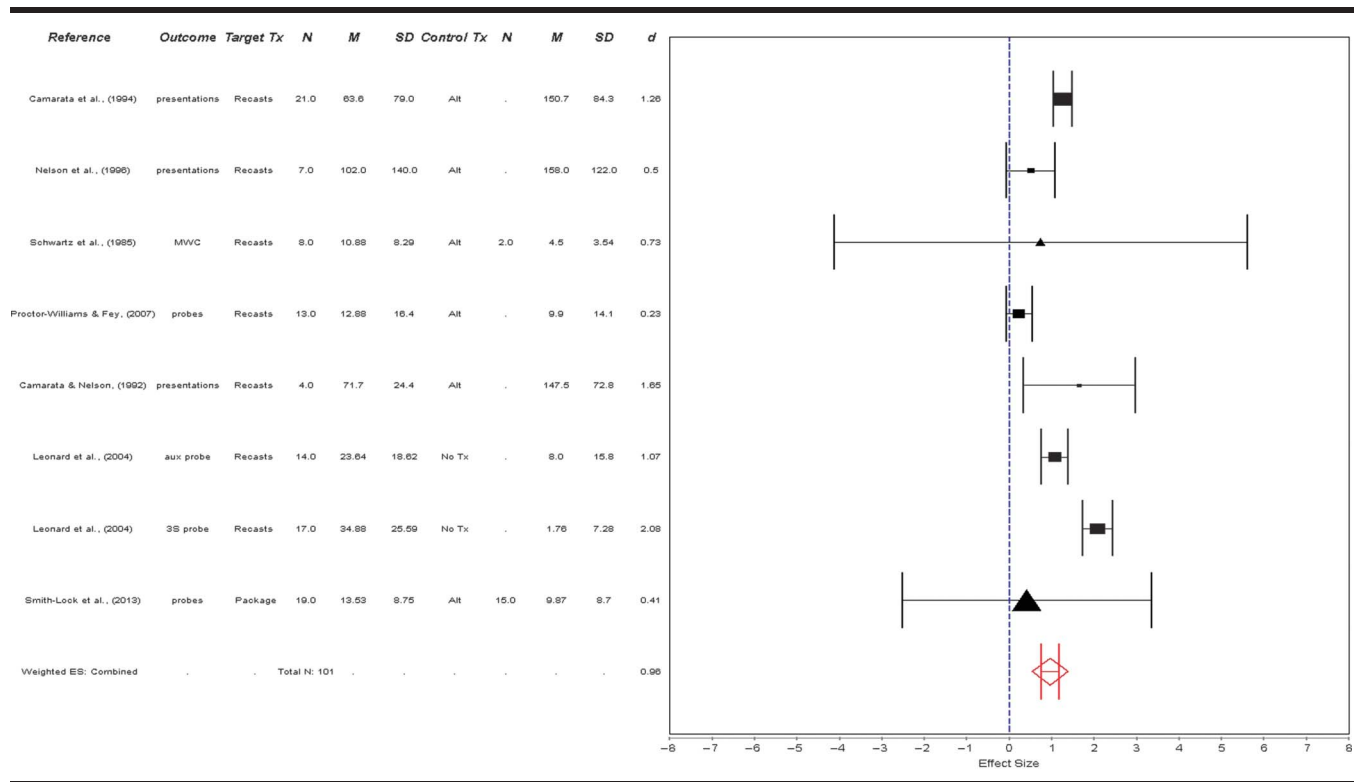
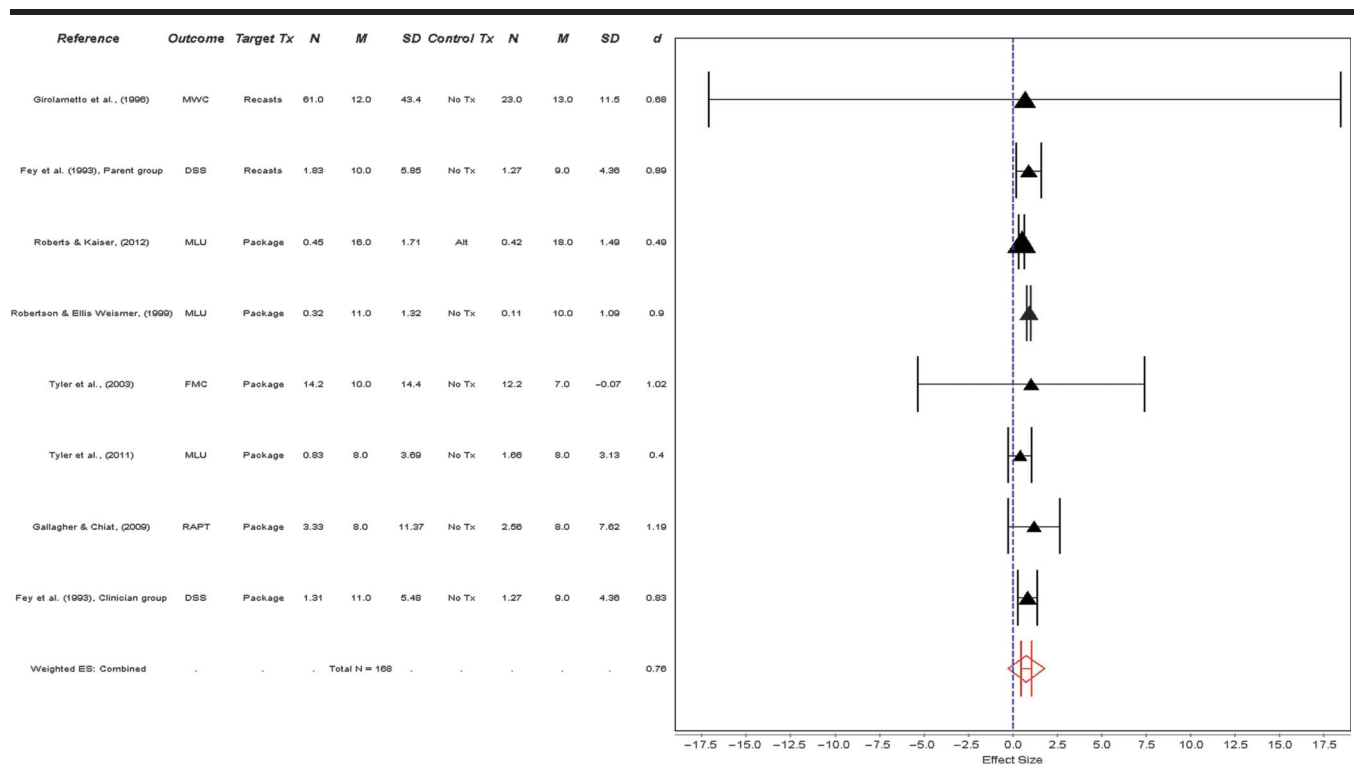


Figure 2. Forest plot of Later Efficacy/Effectiveness studies.



Although it would be tempting to reclassify studies further into “pure” recast versus hybrid approaches, it is difficult to know where to draw the line in terms of grouping studies. While some studies (e.g., Girolametto et al., 1996) clearly used additional treatment techniques and drew on larger packaged interventions, others are less clear. For instance, Leonard et al. (2006) used recasts combined with syntax stories in which target forms were frequently modeled. Is this a pure approach? Additional attention to the active ingredient in an intervention is critical for understanding language treatment effects (Warren, Fey, & Yoder, 2007).

More clear differences between studies include differences in duration and quality. Among the Early Efficacy studies, it is notable that all four of the significant outcomes were found in studies with intervention durations of 12 weeks or longer and quality ratings of 7–9. No significant effects were found for studies with durations of 10 weeks or less, including a study with a quality rating of 9 (Proctor-Williams & Fey, 2007). This trend was not replicated in the Later Efficacy/Effectiveness group, where duration and quality rating do not appear to be consistently associated with differences in effect. The lack of an apparent association between duration or quality rating and outcomes in the Later Efficacy/Effectiveness studies is difficult to interpret and may reflect the variety of packages, procedures, and targets captured in this group. Better reporting of cumulative intensity and dose might be informative for better understanding these findings (Warren et al., 2007).

Question 2: The Impact of Features of the Recasting Intervention

Beyond the basic issue of identifying the evidence for the effects of recasting, we also examined the studies included in our systematic review to determine if different features of recasting intervention resulted in larger or more consistent effects. These principally qualitative comparisons should be viewed as tentative, because the number of studies that speak to an issue is small and often they vary by more than one feature. It is likely, in these cases, that the interaction of features is important. Although we had hoped to address some questions, such as effects of dose or recast rate, through meta-analytic techniques, this was not possible because too few studies were designed to address the relevant comparisons or reported adequate information for moderator analyses to be carried out.

Effects of Intervention Agent

The majority of the studies involved the experimenter or clinician administering the recast intervention; however, there were a few studies in which parents were trained to be the intervention agent. For studies that supported effect-size calculations (Fey et al., 1993; Girolametto et al., 1996), these parent-training programs resulted in positive outcomes with moderate to large effect sizes that were within the range of those seen in the clinician-administered studies. Fey et al. (1993) provide evidence that the parents increased their use of recasting following the training program

compared with parents who did not receive the training. Baxendale and Hesketh (2003) reported positive and equivalent changes in use of language facilitation techniques, including recasts, by parents who attended group training and those who were trained during individual sessions. However, the program that was the basis of the reports of Weistuch and colleagues (Weistuch & Brown, 1987; Weistuch et al., 1991) did not significantly increase recasting by parents. Thus, although parents can be effective intervention agents in recast interventions, it is important to ensure that the training is having an effect on parent behavior. A recent study by Roberts, Kaiser, Wolfe, Bryant, and Spidaleri (2014) investigated a systematic method for training and monitoring parental uses of techniques.

Effects of Type of Recast

Across the studies, there were noteworthy differences in the types of recast used. In most studies, the recasts focused on specific targets, making them part of a focused stimulation approach. In a few studies, however, broad recasts that expanded any child utterance in any way were employed. The only study in which a negative effect size was seen was in one with BTRs (Yoder et al., 2011). As reviewed earlier, that article reported that BTRs were not as effective as MLT for children with MLUs less than 1.84. The other studies that employed BTRs did not allow calculation of effect sizes. However, their results reveal that the use of BTRs did not result in better outcomes (e.g., Loeb & Armstrong, 2001; Yoder et al., 2005) and that positive effects were seen for only some children (e.g., Camarata et al., 2006; Yoder et al., 1995). Furthermore, Leonard et al. (2008) reported that children in the focused stimulation intervention that used focused recasts demonstrated greater gains than those in a general language stimulation that included BTRs, with effect sizes of $d^* = 0.62$, 95% CI [-20.12, 21.78], and $d^* = 0.97$, 95% CI [14.33, 15.97], for third person singular “s” (3s) and auxiliary targets, respectively.

Most studies were not explicit about other recast features or included a mix of recast types. An exception is the article by Hassink and Leonard (2010), which explored the impact of recasting following prompted versus spontaneous child utterances and the impact of corrective versus noncorrective recasts. This study was a follow-up analysis of the treatments reported originally by Leonard et al. (2004, 2006, 2008). Hassink and Leonard report that the frequency of prompted platform utterances was not correlated with the children’s use of the target, suggesting that recasts that followed prompted child utterances were as effective as those that followed spontaneous productions. In addition, the use of noncorrective recasts was positively associated with the children’s use of the target, indicating that recasts do not need to correct a child’s error in order to be effective. However, it should be noted that the noncorrective recasts were still focused recasts because they involved the child’s therapy target. Finally, one of the studies with typically developing children explored the impact when the child’s utterance was recast versus when the

clinician recast her own utterance (Baker & Nelson, 1984). The children who heard recasts of their own utterances exhibited better progress. Additional research into different types of recasts is needed. Although there is a clear pattern in the limited evidence available, suggesting that focused recasts are most effective, any conclusions about the effects of the other recast features must be viewed as tentative.

Effects of Participant Characteristics

Although intervention with recasts results in growth in grammatical ability for children with LI as a group, multiple researchers have noted that there may be differences in participants that influence success in recasting interventions. It is notable that the only study with a negative effect size for recasts (Yoder et al., 2011) identified a conditional treatment effect favoring the use of MLT with children with initially low MLUs, as reviewed earlier. However, studies that involved more focused recasts with children at an early developmental level do report positive effects with recasting (Girolametto et al., 1996; Robertson & Ellis Weismer, 1999; Schwartz, Chapman, Terrell, et al., 1985), suggesting that recasting can facilitate the development of early word combinations.

The studies by Camarata, Nelson, and colleagues (Camarata & Nelson, 1992; Camarata et al., 1994; Nelson et al., 1996) revealed that recasting led to earlier spontaneous production than imitation, even for morphemes not present in production at pretest. With that said, it is possible that children demonstrate better results when a certain level of ability prior to intervention exists in relationship to the target. Leonard and colleagues have reanalyzed their results and identified some possibilities that apply to the particular morphemes they examined. Hassink and Leonard (2010) showed that children's outcomes were influenced by the child's ability to produce subject-verb combinations as platform utterances. In a further analysis of the data from the same studies, Pawlowska, Leonard, Camarata, Brown, and Camarata (2008) found that children's use of plural *-s* at pretest was positively associated with their response to the recast intervention for agreement morphemes. It is unclear how generalizable these findings are to other treatment targets. In contrast to most of the studies reviewed, Fey and Loeb (2002) did not find a positive impact of recasting. Their study focused on auxiliary development, and the target auxiliaries were always recast into yes/no questions in which the auxiliary was inverted. The authors suggest that the lack of effect may have occurred because use of the recast was too advanced relative to the children's language system.

Thus, there is some evidence suggesting that children's developmental level may affect their response to recasts. However, additional research is needed. For instance, the studies by both Hassink and Leonard (2010) and Pawlowska et al. (2008) involved correlational analyses. As interesting and plausible as those findings could be, the level of evidence supporting the claims is weak. For example, although the quality of these studies was high, both were classified as Feasibility studies because of their correlational design. A

useful instantiation of the Fey and Finestack (2009) model for research and development of child language interventions would be for investigators to test the correlational findings of these two studies directly using experimental methods.

The systematic review included five studies involving children with ID. Although it would have been ideal to compare the impact of recasts for a subgroup of studies having children with ID to the larger group of studies of children not having ID, it was not possible to calculate effect sizes due to the lack of an appropriate control comparison (McLean & Vincent, 1984; Weistuch & Brown, 1987) or to insufficient data (Camarata et al., 2006; Weistuch et al., 1991; Yoder et al., 1995). The studies that lacked a control comparison did show growth following recasting, but it is impossible to attribute the growth to recasting. The results of the other three were mixed. Five of the six participants in the study by Camarata et al. (2006) showed growth in MLU, but in only two cases was there evidence of a treatment effect. In a study involving four children, Yoder et al. (1995) found recasting effective for children at the one-word stage but not for those with more developed language. In contrast, in the single-group study (Weistuch et al., 1991), children who received recast intervention showed greater gains than those who received no intervention. In all of the studies, broad recasts were used.

Effects of Dosage

There is reason to believe that recast rate should affect outcomes (Proctor-Williams, 2009; Warren et al., 2007). Analysis of data from individual studies has generally found that differing rates of recasts are associated with differing degrees of progress on proximal outcome measures (Girolametto et al., 1999; Hassink & Leonard, 2010) but that the relationship may be complex and may change as the child becomes more proficient with a target structure (Fey et al., 1993). Evidence suggests that targeting a rate of approximately 0.8–1.0 recast/min may be beneficial (Camarata et al., 1994; Fey et al., 1999). The results from Proctor-Williams and Fey (2007) suggest that too high a rate may reduce efficacy, as has been shown with typically developing children. One might imagine a minimum dose below which no effect is observed (Strain, 2014). It would ideally be possible to analyze recast rates in each study and test whether effect sizes differ for studies with different recast rates and to what degree recast rate must change before a difference in effect size is observed. This type of analysis could similarly be extended to compare time spent in the intervention and total recasts provided over time, to examine the degree to which the intensity within the session or the total recasts over time affects outcomes. Statistical analysis of the type described previously is not possible with our corpus of studies, however, because recast rate is rarely reported or is reported as a target recast rate and not as the actual rate achieved.

Effects of Recast-Only Versus Package Interventions

Yoder et al. (2011) directly compared a recast-only treatment, BTR, to a package, MLT, and found that for

younger children, MLT showed better results. MLT involved imitation, but it also used focused rather than broad recasts. It is impossible to tease apart the impact of these two factors. The meta-analysis did not allow comparison of the effects of recast-only and package interventions because the intervention type covaried with outcomes reported and study design. Recast-only interventions were primarily Early Efficacy studies that reported proximal outcomes and used within-subject designs, whereas package interventions were primarily Late Efficacy/Effectiveness studies that reported distal outcomes and used between-subjects designs. One would predict greater gains in Early Efficacy studies because their outcome measures are tightly tied to the intervention itself (cf. Nye, Foster, & Seaman, 1987). Furthermore, Early Efficacy studies generally involved more intensive, longer term intervention, which would lead to an expectation of greater gains. However, these studies typically compared recast-based treatments to alternate treatments, whereas Late Efficacy studies most often used a no-treatment control, leading to the prediction that Late Efficacy studies should show greater gains. Although Early Efficacy studies had a slightly larger average effect size, it was not significantly larger than that of the Later Efficacy/Effectiveness studies. The range of effect sizes was similar in the two analyses, suggesting similar results.

Because it appears that package interventions do not result in greater gains than recast-only interventions, one might be tempted to conclude that recasts must be the effective ingredient in the package treatments, and other techniques are distractions. However, such a conclusion would be premature. First, the package interventions typically differ not only from recast-only interventions but also from each other. For example, the study by Roberts and Kaiser (2012) is an investigation of the efficacy of parent-implemented enhanced milieu teaching, whereas the study by Smith-Lock, Leita, Lambert, and Nickels (2013) involves an explicit teaching component in a school-based setting. Second, the package interventions included multiple goals across domains. Indeed, considering that children who receive language intervention often exhibit needs across linguistic domains, it is not surprising that several of the package interventions attempted to address more than grammatical skills. For example, the two studies by Tyler and colleagues (Tyler, Gillon, Macrae, & Johnson, 2011; Tyler, Lewis, Haskill, & Tolbert, 2003) consisted of package interventions that targeted grammatical and phonological production for children who demonstrated impairment in both areas. In fact, Tyler et al. (2003) reported that the greatest gains in grammatical skills were seen when the grammar intervention was alternated weekly with the phonological intervention, a finding missed in our meta-analysis because we selected the grammar-only intervention for examination. Certain components of package interventions may be essential to address goals across linguistic domains. Gains in another language domain may enhance gains in grammatical targets. Further work in this area is warranted, and careful reporting of both treatment techniques and the intended gains would enhance follow-up analyses.

The variation between studies utilizing either recast-only or package interventions limits our ability to generate strong conclusions at this time. Studies that differed along carefully selected dimensions and reported comparable outcomes would result in a clearer understanding of the impact of recasts, alone or with specific other techniques. The fact that significant average effect sizes were found for both types of studies is supportive of the use of recasts, but additional studies are needed.

Conclusion

The results of systematic reviews and meta-analyses are most easily applied when the studies included have involved the same intervention with similar participants. We did not find the state of the research on recasts to be so orderly in the study we conducted. Instead, there are many intervention dimensions along which studies vary: comparison group, outcome measure, type of recast employed, recast dosage, contents of intervention packages, age or level of child, and so forth. Furthermore, there are relatively few studies designed to evaluate the contribution of any single feature of a recast intervention.

Given these challenges, there are several modifications of study methodology and reporting practices that researchers could make to facilitate integration of related study outcomes, generally making them more useful to clinicians and more readily applicable in evidence-based decisions. First, preintervention and postintervention data—including number of participants, means, and standard deviations for each participant group—should be reported clearly. When a study uses a within-subject design, individual participant data should be available to enable computation of correlations between pre–post measures, or the authors should present the correlations themselves. Second, in reports of treatment studies of reasonably long duration, both distal and proximal measures should be presented, along with information about the reliability of measures used. Third, a measure of fidelity to intervention protocol, including the actual dosage of the active treatment component, should be provided. For example, a study examining the use of recasts in intervention should report the average number of recasts per minute along with standard deviations for a representative sample of intervention sessions and the actual number of minutes the child spent in therapy. Checklists for following the treatment protocol and intended dose may also be useful information but should not replace actual dose. Fourth, investigators should avoid selecting only children who do not exhibit the target language form, relation, or act. Requiring that study participants be at zero performance levels compromises the ability to calculate meta-analysis statistics in an unbiased way (Schmidt & Hunter, 2015) and makes it more likely that regression to the mean will be observed.

Notwithstanding these limitations in the available research, the evidence from our analyses makes at least two tentative recommendations reasonable. First, our systematic review and meta-analyses support the use of recasts in

programs to facilitate the use of grammatical targets by children with specific LI when they are focused on specific intervention targets and, possibly, complemented by non-contingent models of the target. Clinicians may expect that more general and less concentrated use of recasts will yield outcomes that are not as clinically or statistically significant as those yielded by target-specific recasting. Recommendations for children with ID are more tentative because the results were mixed, and all studies used broad-based recasts. However, although studies examining recasts alone are inconclusive, package approaches that incorporate recasts as a key ingredient do appear to be effective (Roberts & Kaiser, 2011, 2013). Parents and other adults can be effective purveyors of recasts, but clinicians should not assume that all parents will respond to training programs by producing significant increases in parental recast rates outside typical teaching contexts. Parents' actual use of recasts can vary considerably and should be carefully monitored.

Acknowledgments

The meta-analysis work was supported by a grant from the ASH Foundation to Amanda J. Owen Van Horne. Frank L. Schmidt provided advice and support for the within-subject calculations of effect size for the meta-analysis. Paul Yoder and Anne Tyler provided unpublished information on Yoder et al. (2011) and Tyler et al. (2003). Preliminary results from the research were presented at the Annual Convention of the American Speech-Language-Hearing Association in Philadelphia, PA, in November 2010; the Symposium on Research in Child Language Disorders in Madison, WI, in June 2014; and the 2014 DeLTA Day at the University of Iowa.

References

- Baker, N. D., & Nelson, K. E.** (1984). Recasting and related conversational techniques for triggering syntactic advances by young children. *First Language, 5*, 3–22.
- Baxendale, J., & Hesketh, A.** (2003). Comparison of the effectiveness of the Hanen Parent Program and traditional clinic therapy. *International Journal of Language & Communication Disorders, 38*, 397–415.
- Bellon-Harn, M. L.** (2012). Dose frequency: Comparison of language outcomes in preschool children with language impairment. *Child Language Teaching and Therapy, 28*, 225–240.
- Bradshaw, M. L., Hoffman, P. R., & Norris, J. A.** (1998). Efficacy of expansions and cloze procedures in the development of interpretations by preschool children exhibiting delayed language development. *Language, Speech, and Hearing Services in Schools, 29*, 85–95.
- Cable, A. L., & Domsch, C.** (2011). Systematic review of the literature on the treatment of children with late language emergence. *International Journal of Language & Communication Disorders, 46*, 138–154.
- Camarata, S. M., & Nelson, K. E.** (1992). Treatment efficiency as a function of target selection in the remediation of child language disorders. *Clinical Linguistics & Phonetics, 6*, 167–178.
- Camarata, S. M., & Nelson, K. E.** (2006). Conversational recast intervention with preschool and older children. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of language disorders in children* (pp. 237–264). Baltimore, MD: Brookes.
- 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*, 1414–1423.
- Camarata, S. M., 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.
- Cirrin, F. M., & Gillam, R. B.** (2008). Language intervention practices for school-age children with spoken language disorders: A systematic review. *Language, Speech, and Hearing Services in Schools, 39*, S110–S137.
- Cohen, J.** (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York, NY: Academic Press.
- Dollaghan, C.** (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Brookes.
- Ebbels, S.** (2014). Effectiveness of intervention for grammar in school-aged children with primary language impairments: A review of the evidence. *Child Language Teaching and Therapy, 30*, 7–40.
- Farrar, M. J.** (1992). Negative evidence and grammatical morpheme acquisition. *Developmental Psychology, 28*, 90–98.
- Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J.** (1993). *MacArthur–Bates Communicative Development Inventories*. San Diego, CA: Singular.
- Fey, M. E.** (1986). *Language intervention with young children*. Boston, MA: College-Hill.
- Fey, M. E., Cleave, P. L., & Long, S. H.** (1997). Two models of grammar facilitation in children with language impairments: Phase 2. *Journal of Speech, Language, and Hearing Research, 40*, 5–19.
- Fey, M. E., Cleave, P. L., Long, S. H., & Hughes, D. L.** (1993). Two approaches to the facilitation of grammar in children with language impairments: An experimental evaluation. *Journal of Speech and Hearing Research, 36*, 141–157.
- Fey, M. E., & Finestack, L.** (2009). Research and development in child language intervention: A five phase model. In R. Schwartz (Ed.), *Handbook of child language disorders* (pp. 513–531). New York, NY: Psychology Press.
- Fey, M. E., Krulik, T. E., Loeb, D. F., & Proctor-Williams, K.** (1999). Sentence recast use by parents of children with typical language and children with specific language impairment. *American Journal of Speech-Language Pathology, 8*, 273–286.
- Fey, M. E., & Loeb, D. F.** (2002). An evaluation of the facilitative effects of inverted yes-no questions on the acquisition of auxiliary verbs. *Journal of Speech, Language, and Hearing Research, 45*, 160–174.
- Fey, M. E., Richard, G. J., Geffner, D., Kamhi, A. G., Medwetsky, L., Paul, D., . . . Schooling, T.** (2011). Auditory processing disorder and auditory/language interventions: An evidence-based systematic review. *Language, Speech, and Hearing Services in Schools, 42*, 246–264.
- Gallagher, A., & Chiat, S.** (2009). Evaluation of speech and language therapy interventions for pre-school children with specific language impairment: A comparison of outcomes following specialist intensive, nursery-based and no intervention. *International Journal of Language & Communication Disorders, 44*, 616–638.
- Gillum, H., Camarata, S., Nelson, K. E., & Camarata, M.** (2003). A comparison of naturalistic and analog treatment effects in children with expressive language disorder and poor pre-intervention imitation skills. *Journal of Positive Behavior Interventions, 5*, 171–178.

- Girolametto, L., Pearce, P. S., & Weitzman, E.** (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech and Hearing Research, 39*, 1274–1283.
- Girolametto, L., Weitzman, E., Wiigs, E., & Pearce, P. S.** (1999). The relationship between maternal language measures and language development in toddlers with expressive vocabulary delays. *American Journal of Speech-Language Pathology, 8*, 364–374.
- Grela, B. G., & McLaughlin, K. S.** (2006). Focused stimulation for a child with autism spectrum disorder: A treatment study. *Journal of Autism and Developmental Disorders, 36*, 753–756.
- Hancock, T. B., & Kaiser, A. P.** (2006). Enhanced milieu teaching. In R. J. McCauley & M. E. Fey (Eds.), *Treatment of language disorders in children* (pp. 203–236). Baltimore, MD: Brookes.
- Hassink, J., & Leonard, L.** (2010). Within treatment factors as predictors of outcomes following conversational recasting. *American Journal of Speech-Language Pathology, 19*, 213–224.
- Hovell, M., Schumaker, J., & Sherman, J.** (1978). A comparison of parents' models and expansions in promoting children's acquisition of adjectives. *Journal of Child Development, 25*, 41–57.
- Law, J., Garrett, Z., & Nye, C.** (2004). The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 47*, 924–943.
- Leonard, L. B.** (2011). The primacy of priming in grammatical learning and intervention: A tutorial. *Journal of Speech, Language, and Hearing Research, 54*, 608–621.
- Leonard, L. B., Camarata, S. M., Brown, B., & Camarata, M. N.** (2004). Tense and agreement in the speech of children with specific language impairment: Patterns of generalization through intervention. *Journal of Speech, Language, and Hearing Research, 47*, 1363–1379.
- Leonard, L. B., Camarata, S. M., Pawlowska, M., Brown, B., & Camarata, M. N.** (2006). Tense and agreement morphemes in the speech of children with specific language impairment during intervention: Phase 2. *Journal of Speech, Language, and Hearing Research, 49*, 749–770.
- Leonard, L. B., Camarata, S. M., Pawlowska, M., Brown, B., & Camarata, M. N.** (2008). The acquisition of tense and agreement morphemes by children with specific language impairment during intervention: Phase 3. *Journal of Speech, Language, and Hearing Research, 51*, 120–125.
- Loeb, D. F., & Armstrong, N.** (2001). Case studies on the efficacy of expansions and subject-verb-object models in early language intervention. *Child Language Teaching and Therapy, 17*, 35–53.
- McCartney, K., & Rosenthal, R.** (2000). Effect size, practical importance, and social policy for children. *Child Development, 71*, 173–180.
- McLean, M., & Vincent, L.** (1984). The use of expansions as a language intervention technique in the natural environment. *Journal of Early Intervention, 9*, 57–66.
- Nelson, K. E.** (1977). Facilitating children's syntax acquisition. *Developmental Psychology, 13*, 101–107.
- Nelson, K. E.** (1989). Strategies for first language teaching. In M. L. Rice & R. L. Schiefelbusch (Eds.), *The teachability of language* (pp. 263–310). Baltimore, MD: Brookes.
- Nelson, K. E., Camarata, S. M., Welsh, J., Butkovsky, L., & Camarata, M.** (1996). Effects of imitative and conversational recasting treatment on the acquisition of grammar in children with specific language impairment and younger language-normal children. *Journal of Speech and Hearing Research, 39*, 850–859.
- 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*, 497–504.
- Nye, C., Foster, S., & Seaman, D.** (1987). Effectiveness of language intervention with the language/learning disabled. *Journal of Speech and Hearing Disorders, 52*, 348–357.
- Pawlowska, M., Leonard, L., Camarata, S., Brown, B., & Camarata, M.** (2008). Factors accounting for the ability of children with SLI to learn agreement morphemes in intervention. *Journal of Child Language, 35*, 25–53.
- Petersen, D. B.** (2011). A systematic review of narrative-based language interventions with children who have language impairment. *Communication Disorders Quarterly, 32*, 207–220.
- Proctor-Williams, K.** (2009). Dosage and distribution in morpho-syntax intervention. *Topics in Language Disorders, 29*, 294–311.
- Proctor-Williams, K., & Fey, M. E.** (2007). Recast density and acquisition of novel irregular past tense verbs. *Journal of Speech, Language, and Hearing Research, 50*, 1029–1047.
- Proctor-Williams, K., Fey, M. E., & Loeb, D. F.** (2001). Parental recasts and production of copulas and articles by children with specific language impairment and typical language. *American Journal of Speech-Language Pathology, 10*, 155–168.
- Roberts, M. Y., & Kaiser, A. P.** (2011). The effectiveness of parent-implemented language interventions: A meta-analysis. *American Journal of Speech-Language Pathology, 20*, 180–199.
- Roberts, M. Y., & Kaiser, A. P.** (2012). Assessing the effects of a parent-implemented language intervention for children with language impairments using empirical benchmarks: A pilot study. *Journal of Speech, Language, and Hearing Research, 55*, 1655–1670.
- Roberts, M. Y., Kaiser, A. P., Wolfe, C. E., Bryant, J. D., & Spidalieri, A. M.** (2014). Effects of the teach-model-coach-review instructional approach on caregiver use of language support strategies and children's expressive language skills. *Journal of Speech, Language, and Hearing Research, 57*, 1851–1869.
- Robertson, S., & Ellis Weismer, S.** (1999). Effects of treatment on linguistic and social skills in toddlers with delayed language development. *Journal of Speech, Language, and Hearing Research, 42*, 1234–1248.
- Saxton, M.** (1997). The contrast theory of negative input. *Journal of Child Language, 24*, 139–161.
- Saxton, M.** (1998). Longer-term effects of corrective input: An experimental approach. *Journal of Child Language, 25*, 701–721.
- Scherer, N. J., & Olswang, L. B.** (1984). Role of mothers' expansions in stimulating children's language production. *Journal of Speech and Hearing Research, 27*, 387–396.
- Scherer, N. J., & Olswang, L. B.** (1989). Using structured discourse as a language intervention technique with autistic children. *Journal of Speech and Hearing Disorders, 54*, 383–394.
- Schmidt, F. L., & Hunter, J. E.** (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Los Angeles, CA: Sage.
- Schwartz, R. G., Chapman, K., Prelock, P. A., Terrell, B. Y., & Rowan, L. E.** (1985). Facilitation of early syntax through discourse structure. *Journal of Child Language, 12*, 13–25.
- Schwartz, R. G., Chapman, K., Terrell, B. Y., Prelock, P. A., & Rowan, L. E.** (1985). Facilitating word combinations in language-impaired children through discourse structure. *Journal of Speech and Hearing Disorders, 50*, 31–39.
- Smith-Lock, K. M., Leitao, S., Lambert, L., & Nickels, L.** (2013). Effective intervention for expressive grammar in children with specific language impairment. *International Journal of Language & Communication Disorders, 48*, 265–282.
- Smith-Lock, K. M., Leitao, S., Lambert, L., Prior, P., Dunn, A., Cronje, J., . . . Nickels, L.** (2013). Daily or weekly? The role of treatment frequency in the effectiveness of grammar treatment for children with specific language impairment. *International Journal of Speech-Language Pathology, 15*, 255–267.

-
- Strain, P.** (2014, March). *Lessons learned about high fidelity LEAP replication*. Talk presented at ASH Foundation Implementation Science Summit, Carlsbad, CA.
- Turner, H. M., & Bernard, R. M.** (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Sciences and Disorders*, 33, 42–55.
- Tyler, A., Gillon, G., Macrae, T., & Johnson, R.** (2011). Direct and indirect effects of stimulating phoneme awareness vs. other linguistic skills in preschoolers with co-occurring speech and language impairments. *Topics in Language Disorders*, 31, 128–144.
- Tyler, A., Lewis, K., Haskill, A., & Tolbert, L.** (2003). Outcomes of different speech and language goal attach strategies. *Journal of Speech, Language, and Hearing Research*, 46, 1077–1094.
- Warren, S. F., Fey, M. E., & Yoder, P. J.** (2007). Differential treatment intensity research: A missing link to creating optimally effective communication interventions. *Mental Retardation and Developmental Disabilities*, 13, 70–77.
- Weistuch, L., & Brown, B.** (1987). Motherese as therapy: A programme and its dissemination. *Child Language Teaching and Therapy*, 3, 57–71.
- Weistuch, L., Lewis, M., & Sullivan, M. W.** (1991). Project profile: Use of a language interaction intervention in the preschools. *Journal of Early Intervention*, 15, 278–287.
- 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., Molfese, D., & Gardner, E.** (2011). Initial mean length of utterance predicts the relative efficacy of two grammatical treatments in preschoolers with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 54, 1170–1181.
- Yoder, P., 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., Woynaroski, T., Fey, M. E., & Warren, S. F.** (2014). Effects of dose frequency of early communication intervention in young children with and without Down syndrome. *American Journal on Intellectual and Developmental Disabilities*, 119, 17–32.
- Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W.** (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research*, 50, 1562–1592.

Appendix

Calculation of Effect Sizes

Calculation of Effect Sizes for Between-Subjects Designs

Procedures followed those outlined by Turner and Bernard (2006) for calculation of Hedges's g for studies utilizing between-subjects designs. Per Schmidt and Hunter's (2015) recommendation that the similarity across measures be reflected in notation, Hedges's g is identified as d^* throughout this article. The d^* effect size uses a pooled standard deviation measure. The equation is

$$\text{Effect size } d^* = \left[1 - \frac{3}{4N - 9} \right] \left[\frac{\bar{x}_1 - \bar{x}_2}{s} \right],$$

where N is the total number of participants, \bar{x}_1 and \bar{x}_2 are the means of the target and comparison groups, respectively, and s is the pooled standard deviation. Standard error for d^* can be calculated by

$$SE_{d^*} = \sqrt{\frac{1}{n_1} + \frac{1}{n_2} + \frac{d^2}{2(n_1 + n_2)}} * \left(1 - \frac{3}{4(n_1 + n_2) - 9} \right).$$

To compute weighted averages, the effect size of each study was weighted by the total number of participants in that study.

Calculation of Effect Sizes for Within-Subject Designs

Procedures followed those outlined by Schmidt and Hunter (2015) for calculation for studies utilizing within-subject designs. The equation is

$$\text{Effect size } d_{w/in} = \frac{\bar{x}_1 - \bar{x}_2}{(SD_{pooled} / \sqrt{2(1-r)})},$$

where \bar{x}_1 and \bar{x}_2 are the means of the target and comparison groups, respectively, and SD_{pooled} is the standard deviation of the experimental and control structures combined. Best practice would suggest using SD_{pooled} at pretest, but because so many studies selected children who exhibited no use of the target structures, we used posttest values. The correlation between individual participants' results on measurements in the experimental versus the control condition is r . Within-subject calculations for standard error in d are possible using the formula

$$SE_{d_{w/in}} \left[\frac{(N-1)}{(N-3)} \right] * \left[\frac{1}{N} \right] * \left[2(1-r) + \frac{d_{w/in}^2}{2} \right] * \left[\frac{N}{(N-1)} \right].$$

To compute weighted averages, the effect size was weighted by an adjusted N to correct for erroneously large d values. The adjusted N was computed by

$$N_{adj} = 4 \left(1 + d^2 / 8 \right) / S_{ed}^2,$$

where d is the effect size for that study and S_{ed}^2 is the sampling error variance for the study. The weighted average and confidence intervals were then derived via weighted averages on the basis of the sum for N and N_{adj} .
