LSHSS

Clinical Focus

Use of Imitation Training for Targeting Grammar: A Narrative Review

Sarita L. Eisenberg,[a] Shelley L. Bredin-Oja,[b] and Kasey Crumrine[a]

Purpose: The purpose of this narrative review was to examine the evidence for imitation training for targeting grammar in children with developmental language disorder.
Method: Studies investigating imitation training were compiled from two databases—PsycINFO and Linguistics and Language Behavior Abstracts—as well as from reference lists of several books and articles reviewing language therapy.
Results: Twenty-one studies were reviewed. We summarized the demographic and methodological characteristics of the included studies. We identified substantial differences in the implementation of imitation training across the studies and noted a lack of studies considering how those differences might affect therapy outcome. Studies showed a rapid increase in the production of grammatical targets during imitation training activities, higher performance relative to a control group that did not receive any therapy, and generalization to untrained exemplars. However, the studies did not unequivocally link the increased performance to the imitation training therapy. Studies also showed limited generalization to conversational speech. Conversationally based therapy resulted in faster or higher usage in conversational speech, while imitation training resulted in faster immediate gains in usage during therapy trials.
Conclusion: We believe that imitation training can continue to be a useful tool in a speech-language pathologist's toolbox, but with suggested modifications. Specifically, we suggest use of imitation training as a means for rapidly achieving production of previously absent grammatical targets. However, we do not recommend sole or long-term use of imitation training for working on grammar. Further research is needed to refine our use of the procedure.

An important clinical issue for speech-language pathologists (SLPs) is to use therapy approaches that are supported by evidence. Specifically, SLPs want evidence that a particular therapy approach will lead to the largest amount of change in the shortest amount of time. One highly popular procedure used by SLPs is elicited imitation—the provision of a prompt for a child to replicate an utterance modeled by the clinician. In a recent survey by Finestack and Satterlund (2018), 98% of practicing SLPs providing early intervention (i.e., serving children ages 0–5 years) reported using requests for elicited imitation (73% frequently and 25% sometimes), while 0% reported never using this intervention technique. SLPs serving elementary school children responded similarly, with 95% reporting use of elicited imitation (61% frequently and 34% sometimes) and only 2% reporting that they never used this technique. Clearly, SLPs who serve children rely on elicited imitation as a teaching strategy. It is, therefore, incumbent upon our field to examine the evidence for elicited imitation and determine its place in evidence-based practice. It was also reported in the survey that 83% of SLPs in early intervention and 91% of SLPs in elementary education used drill activities either sometimes or frequently. Considering these findings together, it seemed likely that elicited imitation is often used within drill activities. In this report, we specifically examine the evidence for the use of elicited imitation within the context of highly structured drill activities, a procedure termed "imitation training."

Most scholars agree that children do not learn language by imitation alone. That is, imitation as the only learning mechanism is not sufficient to account for the myriad of language skills that a typically developing child acquires. Yet, the ability to imitate has been shown to be an important skill for cognitive and social development (Meltzoff & Williamson, 2013; Tomasello, 1999). Furthermore, children who are more imitative end up with better language. For example, Bates et al. found that an infant's ability to imitate predicted later spoken language ability

[a]Department of Communication Sciences and Disorders, Montclair State University, NJ
[b]Life Span Institute, The University of Kansas, Lawrence
Correspondence to Sarita L. Eisenberg: eisenbergs@montclair.edu

Editor-in-Chief: Holly L. Storkel
Editor: Amanda J. Owen Van Horne
Received May 14, 2018
Revision received August 7, 2019
Accepted November 7, 2019
https://doi.org/10.1044/2019_LSHSS-19-00024
Publisher Note: This article is part of the Forum: Morphosyntax Assessment and Intervention for Children.

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

(Bates et al., 1979, 1989). This effect is seen for older children as well. A study by Speidel and Herreshoff (1989) found that first-grade children who spontaneously imitated more frequently were able to produce longer utterances and that the effects of imitation were significantly more pronounced in the longest category of utterances.

Other researchers have questioned the role of imitation in language learning. In a study of preschool children by Bloom et al. (1974), imitation was limited to structures that children had already acquired or that were produced in the immediately following session. This raises the possibility that imitation might not contribute to language learning if children imitate only things that they already know. Consistent with this possibility and in contrast with Speidel and Herreshoff (1989), Tager-Flusberg and Calkins (1990) found nonimitative utterances to be longer and more advanced than imitative utterances in the spontaneous speech of 2-year-old children.

It must be noted, however, that self-selected or spontaneous imitation—children choosing to imitate an adult's preceding utterance of their own accord—is vastly different from elicited imitation—children imitating an utterance because they have been explicitly instructed to do so. The use of such elicited imitation in therapy has received support in the literature. Fey and Proctor-Williams (2000) recommend the use of elicited imitation as an intervention technique, particularly when first introducing new forms, because it provides practice in production that may strengthen and stabilize syntactic representations. Indeed, imitative responses, when explicitly requested, can be grammatically more advanced than a child's spontaneous utterances (Bredin-Oja & Fey, 2014; Fraser et al., 1963). Additionally, elicited imitation serves to dramatically increase the frequency of producing target structures compared to production rates found in a language learning environment without prompts to imitate and highlights grammatical forms assumed not to be readily perceived by children with developmental language disorder (DLD; Fey & Proctor-Williams, 2000).

Paul and Norbury (2012) note that it may be an advantage that imitation training does not replicate natural language learning. These authors suggest that, because they have not learned language under natural circumstances, children with DLD might need a less natural learning context. The highly structured nature of imitation training enables clinicians to provide condensed input and elicit a large number of responses in a short amount of time, yielding a high dosage of practice trials. Thus, Paul and Norbury recommended using imitation training as a first step in language therapy to establish forms that are absent or infrequently used and then following up with a less structured and more embedded approach to achieve generalization to conversational speech.

The purpose of this review is to examine the evidence for imitation training for targeting grammar in children with DLD. We have reviewed a number of published studies and summarized the, albeit limited and dated, evidence here. Given the quality of studies, we chose to do a narrative review that is broader in scope instead of a systematic review that typically emphasizes randomized controlled studies. The restricted focus of a systematic review has a limited capacity to guide clinical practice because not all the available evidence is considered (MacLure, 2005). Narrative reviews, on the other hand, are comprehensive summaries combined with interpretation and critique and, therefore, can more fully describe the nature of the knowledge base that informs clinical practice (Greenhalgh et al., 2018).

Terminology for Using Models in Therapy

We used the term "model" to refer to any exemplar of a target form presented by the clinician, regardless of whether or not the child is asked to attempt production of the target after hearing the model. While the use of models is an essential component of all language therapies, how models are used varies widely across therapy approaches. The terminology for characterizing the different ways in which models are used in therapy has varied across studies, making it difficult to distinguish between therapy approaches and compare results. Before reviewing the studies, we will, therefore, first suggest a uniform terminology that we will use in our review.

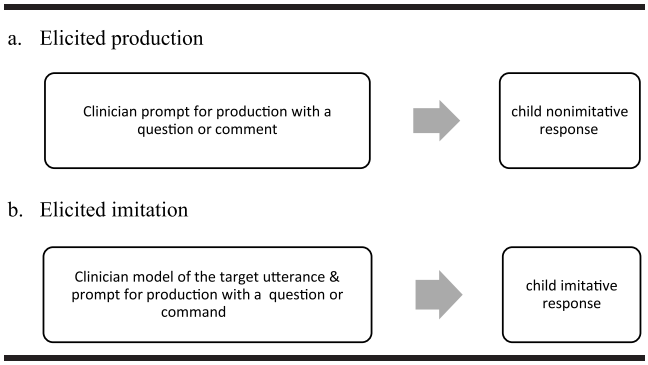
Elicited Production and Imitation

In "elicited production," the clinician uses prompts to encourage the child to attempt the target form. We limited use of this term to instances in which the child is prompted to independently generate the target form. In contrast, we used "elicited imitation" when the child is prompted to replicate a clinician model of the target. Figure 1 illustrates these elicitation procedures. Both elicited production and elicited imitation can be used in the repetitive format of structured drill activities or within embedded communicative interactions to elicit individual responses.

Antecedent and Consequent Models

"Antecedent models" are given before the child's production attempt. That is, the clinician provides a model and prompts the child to replicate that model (i.e., elicited imitation). "Consequent models" are given after the child's production. That is, the clinician prompts the child to generate the target form (i.e., elicited production) and

Figure 1. Comparison of elicited production and elicited imitation.



subsequently models the target. Consequent models that follow incorrect child productions are also referred to as "corrective recasts or expansions" (e.g., Camarata & Nelson, 1992; Goldstein, 1984; Paul & Norbury, 2012). Note though that consequent models can also be given after correct child productions (e.g., Valian & Casey, 2003). After providing a consequent model, the clinician could prompt the child to replicate the modeled utterance (i.e., elicited imitation). Figure 2 illustrates these modeling options.

Imitation Training

We used the term "imitation training" to refer to structured drill activities that involve a series of models and elicited imitations of those models (see Figure 3). Imitation training trials include presentation of a picture or other nonverbal stimulus, a model of the target form, and a prompt to elicit imitation of the model (Eisenberg, 2013; Fey & Proctor-Williams, 2000[1]).

Observational Modeling

Our review includes several studies in which imitation training was compared to a different technique, which we referred to as "observational modeling." In this procedure, the child listens to a set of models of the target form while looking at pictures or watching the clinician use doll figures to act out the utterances (Eisenberg, 2013; Fey & Proctor-Williams, 2000). Observational modeling can be implemented by having the child just listen to the models without subsequently attempting production of the target utterances (e.g., Connell, 1987), or it can be followed by elicited production—prompting the child to talk about the same or a different set of pictures without hearing further models (e.g., Courtright & Courtright, 1976, 1979). Figure 3 illustrates both of these implementation options for observational modeling.

Embedded Therapy

Our review also includes several studies in which imitation training was compared to "embedded therapy" approaches. Such an approach uses contexts in which the child is actively engaged in a communicative interaction rather than a structured drill activity. Within such activities, the clinician provides models of the target form. The clinician might also use instances of either elicited production or elicited imitation to prompt the child to attempt the target form.

Method

PsycINFO and Linguistics and Language Behavior Abstracts were used to search the literature for research studies about imitation training. We used broad search terms to locate reports that used different terminologies to refer to this therapy approach, although this was likely to result in a larger number of irrelevant studies to be manually reviewed. Key word search terms included "child," "language," and "imitation" combined with "intervention," "therapy," or "treatment." We limited the search to peer-reviewed publications written in English. No limits were placed on publication dates, although the search was completed in December 2018. Reference lists from several books and articles reviewing language therapy were also consulted (i.e., Ebbels, 2014; Eisenberg, 2013; Fey, 1986; Leonard, 2014).

The title, abstract, and then the method section of each research study were examined to determine whether the article was a research study that provided evidence about the efficacy of imitation training for targeting grammar goals. This included studies of invented morphemes and syntactic forms. A study was accepted as imitation training if it involved a structured drill activity in which a clinician provided a model and elicited an imitative response from the child. This included studies that compared imitation training to another therapy. The population of interest was children with DLD, but we also accepted studies of children with typical language (TL) as we had no a priori reason to expect that children with TL would respond differently than children with DLD.

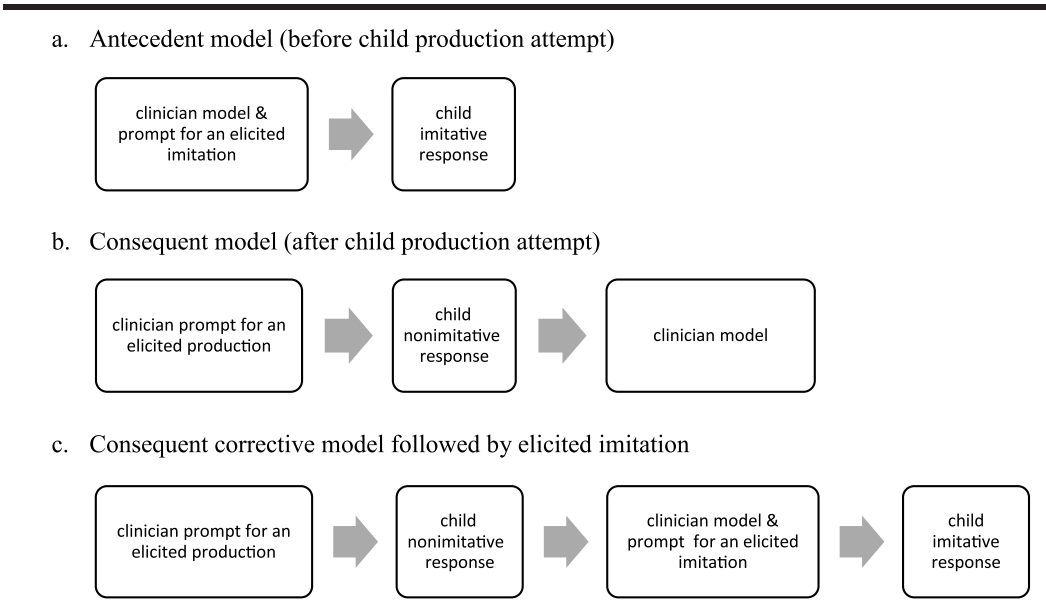
Studies were eliminated from the review for the following reasons: (a) The study was about a topic other than therapy (e.g., assessment, acquisition), (b) the study focused on aspects of language other than grammar (e.g., vocabulary, pragmatics), (c) the study focused exclusively on children with language disorders associated with another syndrome (e.g., autism, intellectual disability, hearing impairment), or (d) the study focused on therapy or instruction for children who were second-language learners. Studies were further restricted based on levels of evidence (LOEs) from the American Academy for Cerebral Palsy and Developmental Medicine (Darrah et al., 2008), listed in Table 1. To be consistent with the broad scope of a narrative review, only studies involving expert opinion and case reports (Level 5) were excluded. All included studies thus had some degree of experimental control.

Results

The database searches yielded 3,400 references after eliminating duplicates. Of these, only 13 studies met the inclusion criteria. An additional eight studies were identified through the reference lists, bringing the total of included studies to 21. Most of the eliminated studies either did not focus on therapy or did not focus on grammar goals. Seventeen studies were excluded because the participants included only children with language disorders associated with autism, intellectual disability, or hearing impairment. Two studies of imitation training were eliminated because they were single-case studies (Level 5). One other group study was eliminated because the therapy did not meet our definition of imitation training. Although a model was provided before each child response, the clinician did not prompt the child to replicate the model.

[1]Note that Fey and Proctor-Williams referred to this as "elicited imitation."

Figure 2. Comparison of antecedent and consequent models.



We first summarized demographic and methodologic characteristics of the included studies. We next reviewed how imitation training was implemented across the studies. Studies were divided into three groups: studies investigating only imitation training, studies comparing imitation training to observational modeling, and studies comparing imitation training to embedded therapies. Within each of these groupings, we separately reviewed studies that used antecedent models and studies that used consequent models.

Demographic and Methodological Characteristics

Source and Publication Date

The included studies were published between 1972 and 2003 in 11 different journals, with 52% of the studies

Figure 3. Comparison of imitation training and observational modeling.

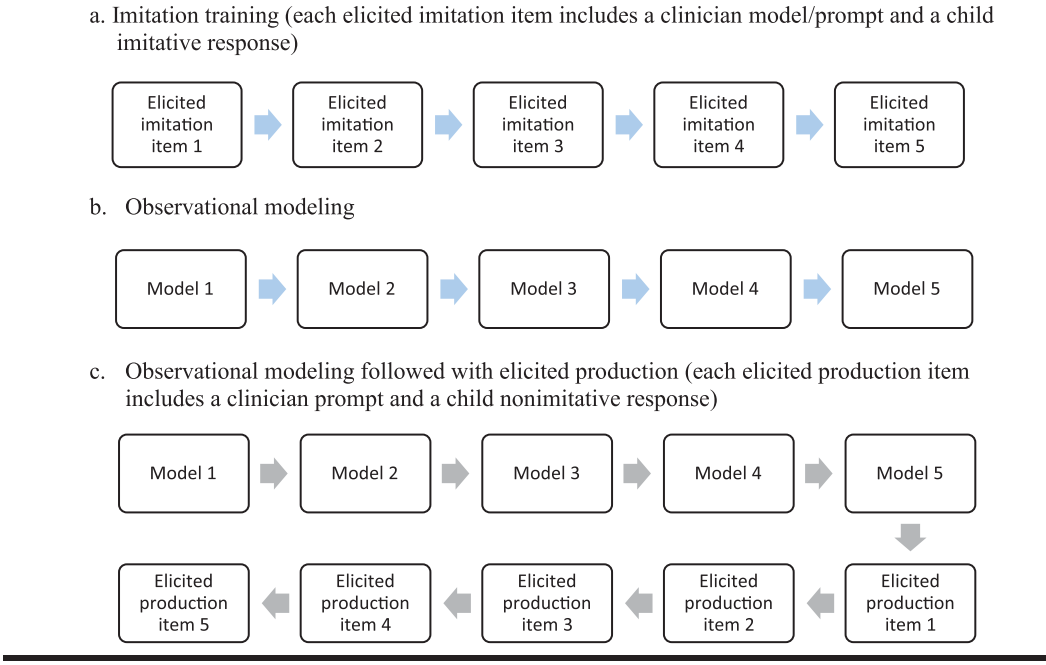


Table 1. Levels of evidence.

Level	Group studies	Single-subject design studies
1	Large randomized[a] controlled trials ($N \geq 100$)	Randomized[b] controlled *N*-of-1 study involving an ATD or a concurrent[c] MBD with staggered initiation of the therapy across ≥ 3 participants or behaviors
2	Smaller randomized[a] controlled trials ($N < 100$)	Nonrandomized, controlled, concurrent[c] MBD across ≥ 3 participants or behaviors
3	Cohort study with concurrent control group	Nonrandomized, nonconcurrent, controlled MBD with ≥ 3 participants or behaviors
4	Cohort study without concurrent control group Case series	Nonrandomized, controlled, single-subject design with ≥ 3 phases (ABA, ABAB, BAB, etc.)
5	Retrospective case reports Expert opinion	Single-case study with AB design

Note. Adapted with permission from the American Academy for Cerebral Palsy and Developmental Medicine (AACPDM; Darrah et al. (2008). ATD = alternating or simultaneous treatment design; MBD = multiple baseline design.

[a]Randomization of participants to condition. [b]Randomization of participants or treatment targets to condition. [c]Concurrent initiation of baseline sessions for participants or targets with staggered initiation of therapy sessions.

published in American Speech-Language-Hearing Association journals (see Table 2). The majority of studies (81%) were published prior to 1995, with a peak between 1975 and 1979 and none published after 2003 (see Figure 4).

Participants

Table 3 shows the number of participants for group and single-subject studies. There were 15 group studies ranging from eight to 80 participants, with most (87%) having fewer than 50 participants and four (27%) having 10 or fewer participants. The six single-subject studies ranged from one to six participants, with three (50%) having only one or two participants. Note that three studies (Camarata & Nelson, 1992; Camarata et al., 1994; Nelson et al., 1996) were classified as group rather than single-subject studies because the data were pooled across participants and mean performance for each condition was reported as for a group design.

Participant age (see Tables 6–8) ranged from 2;2 to 10;0 (years;months), with all but one study including participants between the ages of 3;0 and 6;10. Seventeen studies included only children with DLD, two studies included only children with TL, and four studies included both children with DLD and TL.

Table 2. Sources of included studies.

Journal	No. studies
Applied Psycholinguistics	1
Child Language Teaching and Therapy	1
Clinical Linguistics & Phonetics	1
International Journal of Language & Communication Disorders	1
Journal of Child Language	1
Journal of Communication Disorders	4
Journal of Positive Behavior Interventions	1
Journal of Speech and Hearing Disorders[a]	2
Journal of Speech and Hearing Research[a]	7
Journal of Speech, Language, and Hearing Research[a]	1
Language, Speech, and Hearing Services in Schools[a]	1

[a]A journal published by the American Speech-Language-Hearing Association.

Therapy Targets

Therapy targets are included in Tables 6–8. Most of the studies (73%) targeted the same forms for all participants, including sentence constituents (2), subject case pronouns (2), possessive morpheme *'s* (1), use of auxiliary and/or copula *BE* (5), and questions with auxiliary and/or copula *BE* (2). In five of these studies, the targets were invented grammatical forms, including noun suffixes (2), verb suffixes (1), and word ordering rules (2). One study assigned different targets to each participant—either copula *BE* or auxiliary *BE*. Five other studies individualized targets based on language sample analysis and probes and included noun and verb morphemes, sentence types, and complex sentence forms as the therapy targets.

In 12 (57%) of the studies, all targets were completely absent (i.e., 0% usage) at baseline. One study included only partially mastered targets (i.e., up to 65% usage) at baseline. Five (24%) of the studies included both absent and partially mastered targets. The remaining three studies did not specify the baseline level of the target forms.

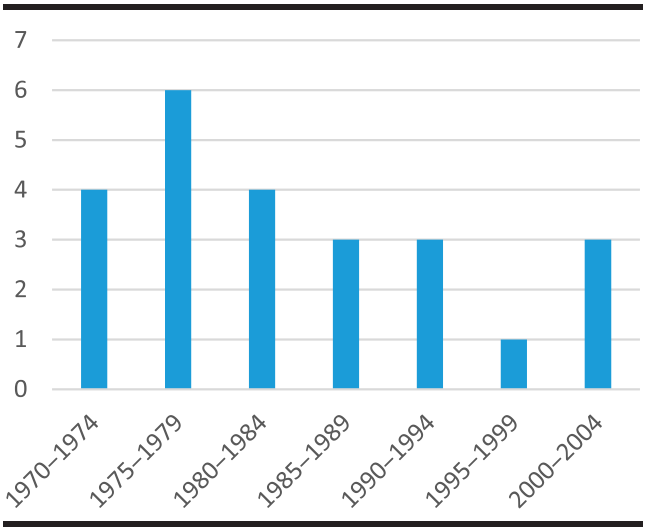
Service Delivery

Service delivery is shown in Table 4. In most (90%) of the studies, therapy was provided individually in a clinic. As can be seen from the table, scheduling and dosage varied considerably.

LOE

Table 5 summarizes the LOE across the included studies. Two thirds of the group studies were classified as Level 2. However, although meeting the randomization criteria for this level, six of the studies were very small. Three of these studies, namely, Camarata and Nelson (1992), Haynes and Haynes (1980), and Mulac and Tomlinson (1977), included 10 or fewer total participants, and three other studies, namely,

Figure 4. Publication dates of the included studies.



Anderson (2001), Courtright and Courtright (1979), and Matheny and Panagos (1978), had group sizes of less than 10. The remainder of the group studies (33%) was classified as Level 3 cohort studies without randomized assignment of participants to condition. Most of the single-subject studies (83%) were classified as either Level 1 or 2. However, although meeting the design criteria for these levels, two of the studies, namely, Hegde et al. (1979) and Hughes and Carpenter (1989), included only one or two participants. Given these limitations on sample size, only seven of the included studies would have met the criteria for a systematic review. The classification of each study is included in Tables 6–8.

Implementation of Imitation Training

Imitation training involves the provision of clinician models and elicited imitation of those models within a highly structured drill activity. All of the studies involved an exact match between the model and the child's response. That is, the clinician provided a model of the target form, and the child was asked to repeat the model verbatim (exact imitation, also called "mimicry"; Haynes & Haynes, 1980). There was otherwise considerable variation across the studies in how imitation training was implemented. We summarize the differences here. Additional details are shown in Table 9.

Table 3. Number of participants in the included studies.

Number of participants	Number (%) of studies
Group design studies	
≤ 10	4 (27)
11–19	1 (7)
20–29	5 (33)
30–39	1 (7)
40–49	2 (13)
≥ 50	2 (13)
Total	15
Single-subject design studies	
1–2	3 (50)
3–6	3 (50)
Total	6

1. Timing of the prompt relative to the child's response: Figure 5 illustrates this variable. In most of the reviewed studies, the clinician provided a prompt before saying the model so that the child's response occurred immediately after the model (immediate imitation, e.g., *Say the boy is running*; Connell, 1987; Connell & Stone, 1992; Hegde, 1980; Hegde & Gierut, 1979; Hegde et al., 1979; Hughes & Carpenter, 1989; Matheny & Panagos, 1978; Mulac & Tomlinson, 1977; Nelson et al., 1996; Valian & Casey, 2003). Alternatively, in some of the studies, the clinician first modeled the target and then prompted the child to imitate (delayed imitation, e.g., *The boy is running. You say that*; Anderson, 2001; Courtright & Courtright, 1976, 1979; Gillum et al., 2003). Timing of the prompt was not specified in the other seven studies.
2. Timing of the model relative to the child's response (see Figure 2): In most of the reviewed studies, models were provided before child responses (antecedent models; Anderson, 2001; Camarata & Nelson, 1992; Camarata et al., 1994; Cole & Dale, 1986; Connell, 1987; Connell & Stone, 1992; Courtright & Courtright, 1976, 1979; Friedman & Friedman, 1980; Gillum et al., 2003; Haynes & Haynes, 1980; Hegde, 1980; Hegde & Gierut, 1979; Hegde et al., 1979; Hughes & Carpenter, 1989; Matheny & Panagos, 1978; Mulac & Tomlinson, 1977; Nelson et al., 1996; Valian & Casey, 2003; Zwitman & Sonderman, 1979). Alternatively, in two of the studies, models were provided after child responses (consequent models; Goldstein, 1984; Valian & Casey, 2003).
3. Contingency of models: For most of the reviewed studies with antecedent models, the child heard a model before each attempt regardless of response accuracy (Anderson, 2001; Camarata & Nelson, 1992; Camarata et al., 1994; Cole & Dale, 1986; Connell, 1987; Connell & Stone, 1992; Courtright & Courtright, 1976, 1979; Friedman & Friedman, 1980; Gillum et al., 2003; Haynes & Haynes, 1980; Hughes & Carpenter, 1989; Matheny & Panagos, 1978; Mulac & Tomlinson, 1977; Nelson et al., 1996; Valian & Casey, 2003; Zwitman & Sonderman, 1979). In some of these studies, the provision of antecedent models was abruptly discontinued after the child achieved some accuracy criterion such that therapy shifted from elicited imitation to elicited production (Camarata & Nelson, 1992; Camarata et al., 1994; Hughes & Carpenter, 1989; Nelson et al., 1996). In several studies, however, antecedent models were contingent on the child's accuracy. Hegde and colleagues discontinued antecedent models after four consecutive correct responses and then reinstated antecedent models for four subsequent

Table 4. Service delivery in the included studies.

General	*n* (%)	Session length[a]	*n* (%)	Sessions per week[a,b]	*n*[b] (%)	Total sessions	*n* (%)
Individual	19 (90)	10–19 min	1 (5)	1	1 (5)	1	1 (5)
Small group	2 (9)	20–25 min	6 (29)	2	5 (25)	3–6	7 (33)
		30–35 min	3 (14)	3	3 (15)	15–18	4 (19)
		40–45 min	3 (14)	4	2 (10)	24–32	3 (14)
		50 min	1 (5)	5	1 (5)	50	1 (5)
		60 min	1 (5)	NS	8 (40)	≥ 125	2 (10)
		120 min	1 (5)			CNT	3 (14)
		NS	5 (24)				

Note. NS = not specified; CNT = could not tell.

[a]The lower value was used for studies with variable session length or sessions per week. [b]The total number sums to only 20 because one study involved a single session.

trials after the child made an error on a nonimitative production (Hegde, 1980; Hegde & Gierut, 1979; Hegde et al., 1979). In one study with only consequent models, a model was provided after each incorrect response, and the child was prompted to repeat the model (termed a "corrective model"; Goldstein, 1984). These options are illustrated in Figure 6.

4. Density of models: In most studies, models were provided with a consistently high density, that is, before every attempt by the child (Camarata & Nelson, 1992; Camarata et al., 1994; Cole & Dale, 1986; Courtright & Courtright, 1976, 1979; Friedman & Friedman, 1980; Gillum et al., 2003; Haynes & Haynes, 1980; Hughes & Carpenter, 1989; Matheny & Panagos, 1978; Mulac & Tomlinson, 1977; Nelson et al., 1996; Valian & Casey, 2003; Zwitman & Sonderman, 1979). A few studies provided an even higher density of models, either by providing two antecedent models of the target sentence before each attempt (Connell, 1987; Connell & Stone, 1992; Nelson et al., 1996; Valian & Casey, 2003) or by also providing a consequent model after the child's attempt, regardless of correctness (Valian & Casey, 2003). An alternative would be to present models less frequently, for instance, before every two or three child responses, so that imitative attempts alternated with nonimitative trials. This occurred in the study by Anderson (2001) because they alternated therapy trials with test trials that were not preceded by a model. For still other studies in which models were provided only when the child made errors, density varied as a function of child accuracy (Goldstein, 1984; Hegde, 1980; Hegde & Gierut, 1979; Hegde et al., 1979). This also resulted in a lower overall density of models.
5. Consequences after attempts: Most of the studies provided some form of reinforcement after correct imitations. The most common reinforcer was praise or social reinforcement (e.g., "good"). In many of the other studies, this was combined with either a primary reinforcer (food) or tokens.

Table 5. Level of evidence of the included studies.

Level of evidence	Number (%) of studies		
	Group	Single subject	All
1	0	2 (33)	2 (10)
2	10 (67)	2 (33)	12 (57)
3	5 (33)	0	5 (23)
4	0	2 (33)	2 (10)
Total	15	6	21

We found only one study that compared different ways of implementing imitation training. The study by Valian and Casey (2003) compared outcome for imitation training with differing density and/or timing of models. This study included three conditions: one in which imitation training was implemented with a single antecedent model before child attempts (see Figure 7a); a second condition in which the child heard two antecedent models before each attempt (see Figure 7b); and a third condition with both an antecedent model and a consequent model, with the latter provided after the child's production regardless of the correctness of the child's imitation (see Figure 7c). Inversion for trained auxiliaries (*BE, can*) increased in all three conditions. However, generalization of inversion to untrained auxiliaries (*will, DO*) occurred only for the two conditions in which the child heard two models, with no difference between these two conditions. We otherwise lack studies comparing these differences in implementation.

Studies Examining the Efficacy of Imitation Training

Three group studies (Matheny & Panagos, 1978; Mulac & Tomlinson, 1977; Zwitman & Sonderman, 1979) investigated the efficacy of imitation training with antecedent models (as illustrated in Figure 2a). Two of these studies reported higher performance after imitation training compared to a no-therapy control condition (Matheny & Panagos, 1978; Zwitman & Sonderman, 1979). However, Matheny and Panagos (1978) also included a control group receiving articulation therapy and that group showed almost as much change in performance on a sentence imitation

Table 6. Studies of imitation training (IT).

Study	Participants	Therapy targets	Baseline	Therapy	Delivery and dose	Level of evidence	Results	Maintenance	Generalize to untrained exemplars	Generalize to speech
Hegde (1980)	DLD 5;6–5;10 (years; months) #IT = 2	S1: aux *is* S2: cop *is*	Absent	Immediate exact imitation	1:1 3 days/wk, 40–45 min	Level 4 SSD	S1 (aux): A1 = increase to 90% usage; B = drop to 30% usage; A2 = increase to 100% usage S2 (cop): A1 = increase to 96% usage; B = drop to 23% usage; A2 = increase to 100%	Not measured	Yes Usage also increased on untrained target: S1 to cop; S2 to aux	Yes
Hegde et al. (1979)	DLD 3;9–4;0 #IT = 2	aux *'s*, cop *'s*, aux *was*, cop *was*, poss *–s*	Absent	Immediate exact imitation	1:1 4–5 days/wk, 45 min	Level 2 SSD	S1 (aux 1st): no change in baseline until therapy began for aux *'s*, aux *was*, and poss S2 (cop 1st): no change in baseline until therapy began for cop *'s*, cop *was*, and poss	Not measured	Yes Usage also increased on untrained target: For S1, to cop after working on aux For S2, to aux after working on cop	Not measured
Hegde & Gierut (1979)	DLD 4;9 #IT = 1	Subject pronouns (*he, she, they*) and aux *are*	Absent	Imitation	1:1; 17 30-min sessions for 4 wks	Level 2 SSD	Reduction in number of trials to achieve criterion (10 correct) on each target sentence	Not measured	Yes	Not measured
Matheny & Panagos (1978)	DLD 5;5–6;10 #IT = 8 #cN = 8 #cA = 8	Not specified	Not specified	Immediate exact imitation	1:1; 5 months	Level 2 group	IT > cA for syntax, but both made significant gains in syntax; cN made no significant gains in syntax	Not measured	Not measured	Not measured

(*table continues*)

Table 6. *(Continued).*

Study	Participants	Therapy targets	Baseline	Therapy	Delivery and dose	Level of evidence	Results	Maintenance	Generalize to untrained exemplars	Generalize to speech
Mulac & Tomlinson (1977)	DLD 4;4–6;3 #IT = 3 #IT+ = 3 #cA = 3	cop and aux *is* questions	Absent	Immediate exact imitation; IT+ added transfer training	1:1; 120 min/wk (60 min, 2 days/wk; 30 min, 4 days/wk)	Level 2 group	Both IT and IT+ achieved 100% usage in therapy sessions; did not report outcome for cA but not significantly different from outcome for therapy groups	Yes, 3 wks after IT+ training concluded	Not measured	No, for IT Yes, for IT+
Valian & Casey (2003)	DLD 2;6–3;2 #IT1 = 11 #IT2 = 9 #IT3 = 9	*wh*-questions	Partially mastered	Exact imitation	1:1; 4 days/wk	Level 3 group	Increase target use in all three conditions; IT1 < IT2 and IT3	Not measured	No for IT1; yes for IT2 and IT3 with IT2 = IT3	Not measured
Zwitman & Sonderman (1979)	DLD 3;4–4;6 #IT = 11 #cN = 11	N-is-Ving-N	Emerging/ partially mastered	Exact imitation	1:1; 1 day/ wk, 45 min	Level 3 group	IT > cN for mean gain in usage of targets, but five in cN made substantial gains	Not measured	Yes	Yes

Note. In baseline, absent = 0% usage, emerging = ≤ 10% usage, and partially mastered = 11%–65% usage. DLD = developmental language disorder; aux = auxiliary; cop = copula; 1:1 = one-to-one individual therapy; wk(s) = week(s); SSD = single-subject design; poss = possessive; cN = no therapy control; cA = articulation therapy control; IT+ = IT with extended transfer training; IT1 = one antecedent model; IT2 = two antecedent models; IT3 = one antecedent model and one consequent model.

Table 7. Studies comparing imitation training (IT) and observational modeling (OM).

a: Studies of imitation training with antecedent models

Study	Participants	Therapy targets	Baseline	Comparison therapy	Delivery and dose	Level of evidence	Results	Maintenance	Generalize to untrained exemplars	Generalize to connected speech
Anderson (2001)	DLD and TL 3;8–6;9 (years; months) #IT = 16 #OM = 16	Invented verb suffix	Absent	Production of nontarget utterance after each model	1:1; one session with 32 trials each	Level 2 group	IT = OM for both DLD and TL	Not measured	Yes	Not measured
Connell (1987)	DLD and TL 3;8–6;0 #IT = 40 #OM = 40	Invented noun suffix	Absent	Just listening	1:1; three 20-min sessions with 20 trials each	Level 2 group	DLD: IT > OM for correct responses to probes; TL: IT < OM for correct responses to probes	Not measured	Yes	Not measured
Connell & Stone (1992)	DLD and TL 3;0–6;11 #IT = 37 #OM = 37	Invented noun suffixes	Absent	Production of nontarget utterance after each model	1:1; four sessions in total, two sessions of 20 trials each	Level 3 group	TL: IT = OM DLD: IT > OM for rate of correct responses to probes	Not measured	Yes	Not measured
Courtright & Courtright (1976)	DLD 5–10 years #IT = 4 #OM = 4	Subject pronoun *they*	IT: emerging/partially mastered OM: absent	Attempted same utterances after hearing sets of 10 utterances	1:1; three 20-min sessions with 20 trials each	Level 3 group	Performance increased in both IT and OM conditions First probe: IT > OM Last probe: IT < OM IT: plateau after second session OM: continued increase across all sessions	Higher accuracy after 1 week in modeling condition	IT < OM (i.e., higher accuracy in OM on probe for untrained exemplars)	Not measured

(*table continues*)

Table 7. *(Continued).*

a: Studies of imitation training with antecedent models

Study	Participants	Therapy targets	Baseline	Comparison therapy	Delivery and dose	Level of evidence	Results	Maintenance	Generalize to untrained exemplars	Generalize to connected speech
Courtright & Courtright (1979)	DLD 3;11–6;11 #IT = 12 #OM = 12	Invented grammatical form	Absent	Attempted novel utterances after hearing sets of 10 utterances	1:1; three 20-min sessions with 20 trials each	Level 2 group	IT < OM for accuracy on probe after each session	Higher accuracy after 1 week in modeling condition	IT < OM (i.e., higher accuracy in OM on probe for untrained exemplars)	Not measured
Haynes & Haynes (1980)	DLD 4 years #IT = 5 #OM = 5	Copula *is*	Absent/ emerging	Attempted novel utterances after hearing sets of 10 utterances	1:1; five sessions with 40 trials each	Level 2 group	IT > OM for rate of correct responding for all five sessions	No difference on probe after 1 week	IT = OM (i.e., not significantly different overall); however, over the four probes, plateau for IT versus consistent increase and higher final level for OM	Not measured

b: Studies of imitation training with consequent models

Goldstein (1984)	TL 4;11–5;7 #IT = 6 Low SES	Nonsense four- to six-word combinations (of previously learned two-word combinations)	Absent	Just listening	1:1; 3–5 days/ week for 20–25 min	Level 1 SSD	IT > OM for all six children Two children learned only in IT (no change in OM)	Not measured	Yes	Not measured

Note. In baseline, absent = 0% usage, emerging = ≤ 10% usage, and partially mastered = 11%–65% usage. DLD = developmental language disorder; TL = typical language; SES = socioeconomic status; 1:1 = one-to-one individual therapy; SSD = single-subject design.

Table 8. Studies comparing imitation training (IT) and conversational therapy (C).

Study	Participants	Therapy targets	Baseline	Comparison therapy	Delivery and dose	Level of evidence	Results	Maintenance	Generalize to untrained exemplars	Generalize to connected speech
Camarata & Nelson (1992)	DLD 4;9–5;11 (years; months) *n* = 4	Varied morphemes and syntactic forms	Absent	CRT	1:1 30 min, 2 days/wk for 32 days	Level 2 group	IT = C for number of presentations prior to achieve first elicited production; IT > C for number of presentations to achieve first spontaneous use	Not measured	Not measured	IT < C for mean number of nontargeted productions
Camarata et al. (1994)	DLD 4;0–6;10 *n* = 21	Varied morphemes and syntactic forms	Absent	CRT	1:1 25 min, 2 days/wk for 24 days	Level 2 group	IT < C for number of clinician presentations and number of sessions to achieve first elicited production; IT > C for number of clinician presentations and number of sessions to achieve first spontaneous production	Not measured	Not measured	IT < C for total number of nontargeted productions

Cole & Dale (1986)	DLD
	3;2–5;9
	#IT = 19
	#C = 25
	IQ = 59–109
		Not specified	Not specified	"Interactive language instruction" similar to CRT	4:1
					2-hr classes, 2 days/wk for IT and 3 days/wk for C for 32 wks
						Level 2 group	IT = C for DSS, PLS, and NSST
							Neither IQ nor language score predicted gain from either therapy
								Not measured	Not measured	Not measured

Friedman & Friedman (1980)	DLD 3;2–5;9 #IT = 16, #C = 25 IQ varied Pretherapy DSS varied	Varied morphemes and syntactic forms	Not specified	ILDT	2:1 in class 1 hr, 4 days/wk for 8 mos	Level 3 group	For gains in DSS: Higher IQ: IT < C Lower IQ: IT > C LL: IT < C, with large gains for IT and moderate gains for C HL: IT < C, with little gains for IT and moderate gains for C	Not measured	Not measured	Yes

(*table continues*)

Table 8. *(Continued).*

Study	Participants	Therapy targets	Baseline	Comparison therapy	Delivery and dose	Level of evidence	Results	Maintenance	Generalize to untrained exemplars	Generalize to connected speech
Gillum et al. (2003)	DLD 4;3–6;8 *n* = 4 Poor imitation skills	Varied morphemes and syntactic forms	Absent	CRT	1:1 25 min, 2 days/wk for 24 days	Level 1 SSD	IT < C for number of trials to criterion (criterion = 5 elicited productions)	Not measured	Not measured	Not measured
Hughes & Carpenter (1989)	4;4–6;3 *n* = 4	Copula *is* and *are*	Absent/ partially mastered	ILDT	1:1 30 min, 3 days/wk for 4 wks	Level 4 SSD	IT > C for three of the participants IT > C for number of targets reaching 80% usage on therapy trials: four for IT and three for C	Not measured	IT = C	Little or no generalization to at-home conversation for either condition
Nelson et al. (1996)	DLD and TL 2;2–6;7 #DLD = 7 #TL = 7	Varied morphemes and syntactic forms; six per child	Three absent and three partially mastered	CRT	1:1 2 days/wk for an average of 18.8 sessions	Level 2 group	Absent targets: IT > C for number of presentations to achieve first spontaneous use for both DLD and TL; three with DLD did not achieve spontaneous production in IT Partially mastered: IT < C for gain in number of spontaneous productions	Not measured	Not measured	IT < C for number of targets generalized to home use (31% vs. 77%)

Note. In baseline, absent = 0% usage and partially mastered = up to 65% usage. DLD = developmental language disorder; CRT = conversational recast therapy (Baker & Nelson, 1984); 1:1 = one-to-one individual therapy; wk(s) = week(s); 4:1 = small group with four children; DSS = Developmental Sentence Score; PLS = Preschool Language Scale; NSST = Northwestern Screening Syntax Test; ILDT = interactive language development teaching (Lee, 1974); mos = months; LL = low language; HL = high language; SSD = single-subject design; TL = typical language.

Table 9. Implementation of imitation training.

Study	Match between response and model	Timing of response after prompt	Timing of the model	Contingency of models	Density of models	Consequence after attempts
Anderson (2001)	Exact	Delayed	Ant	NA	1:1 (before every attempt); test trials alternated with therapy trials	Not specified
Camarata & Nelson (1992)	Exact	Not specified	Ant	Model discontinued after 90% accuracy	1:1 (before every attempt)	c: praise and token reinforcement i: feedback
Camarata et al. (1994)	Exact	Not specified	Ant	Model discontinued after 90% accuracy	1:1 (before every attempt)	c: praise and/or token reinforcement
Cole & Dale (1986)	Exact	Not specified	Ant	NA	1:1 (before every attempt)	c: praise
Connell (1987)	Exact	Immediate	Ant	NA	2:1 (two models before every attempt)	c: praise i: correction request
Connell & Stone (1992)	Exact	Immediate	Ant	NA	2:1 (two models before every attempt)	c: consequating event (an animated character acts out the child's utterance)
Courtright & Courtright (1976)	Exact	Delayed	Ant	NA	1:1 (before every attempt)	No reinforcement
Courtright & Courtright (1979)	Exact	Delayed	Ant	NA	1:1 (before every attempt)	No reinforcement
Friedman & Friedman (1980)	Exact	Not specified	Ant	Contingency for fading the model not specified	1:1 (before every attempt)	c: not specified
Gillum et al. (2003)	Exact	Delayed	Ant	Model discontinued after 90% accuracy	1:1 (before every attempt)	c: praise and/or token reinforcement i: feedback
Goldstein (1984)	Exact	Not specified	Cons	Model after each incorrect response before the second attempt	Varied based on child's accuracy	c: token i: model for imitation
Haynes & Haynes (1980)	Exact	Not specified	Ant	NA	1:1 (before every attempt)	c: social reinforcement i: "no" and another model
Hegde (1980)	Exact	Immediate	Ant	Model discontinued after four correct and reinstated after two incorrect	Varied based on child's accuracy	c: candy and verbal reinforcement i: "no" and averting eye contact
Hegde & Gierut (1979)	Exact	Immediate	Ant	Models initiated on the next trial after error; discontinued after four correct imitative trials	Varied based on child's accuracy	c: candy and verbal reinforcement i: "no" and averting eye contact
Hegde et al. (1979)	Exact	Immediate	Ant	Model discontinued after four correct and reinstated after two incorrect	Varied based on child's accuracy	c: verbal reinforcement and tokens that could be traded for play time i: "no"
Hughes & Carpenter (1989)	Exact	Immediate	Ant	Model discontinued after three sets of five correct responses	1:1 (before every attempt)	c: token and verbal praise i: model for imitation; correction prompt
Matheny & Panagos (1978)	Exact	Immediate	Ant	NA	1:1 (before every attempt); test trials alternated with therapy trials	c: not specified
Mulac & Tomlinson (1977)	Exact	Immediate	Ant	NA	1:1 (before every attempt)	c: not specified

(*table continues*)

Table 9. *(Continued).*

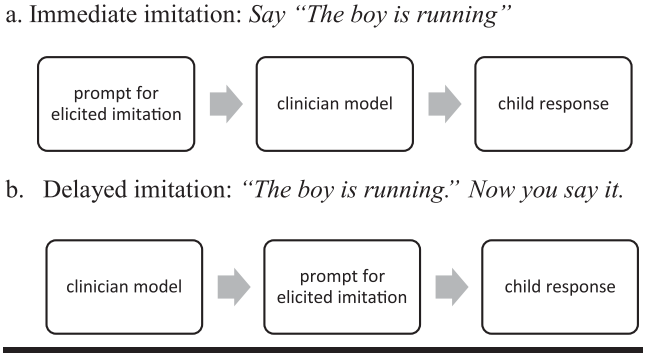
Study	Match between response and model	Timing of response after prompt	Timing of the model	Contingency of models	Density of models	Consequence after attempts
Nelson et al. (1996)	Exact	Immediate	Ant	Model discontinued after 90% accuracy	Two models before every attempt	c: verbal praise and tokens that could be traded for a prize
Valian & Casey (2003)	Exact	Immediate	C1: Ant C2: 2 Ant C3: Ant and Cons	NA	C1: 1:1 (before every attempt) C2: 2:1 (two models before every attempt) C3: 2:1 (two models, one before and one after every attempt)	c: consequating event (toy bear answered the question)
Zwitman & Sonderman (1979)	Exact	Not specified	Ant	Model discontinued after 10 consecutive correct	1:1 (before every attempt)	c: "good" and tokens that could be exchanged for a gift

Note. Ant = antecedent; NA = not applicable; c = correct; i = incorrect; Cons = consequent; C1/C2/C3 = Condition 1/2/3.

task as the imitation group. Similarly, in the study by Mulac and Tomlinson (1977), the gain in performance for the imitation training condition was not significantly different from the gain made by a control group of children who had received articulation therapy. In addition, although Zwitman and Sonderman (1979) reported larger gains in mean production accuracy for the imitation training group, five children in the no-therapy control group showed as large an increase in production as children who had received imitation training.

Several single-subject design studies by Hegde and colleagues (Hegde, 1980; Hegde & Gierut, 1979; Hegde et al., 1979) examined the efficacy of imitation training using contingent antecedent models after errors (illustrated in Figure 6c). Hegde and Gierut (1979) reported an increase in production for untrained and trained exemplars of the target form for each of the two participants. Hegde et al. (1979) conducted a single-subject study with staggered multiple baselines across four or five morpheme targets for two children. Subject 1 showed no change from baseline until after the initiation of therapy for auxiliary *is,* auxiliary *was,* and possessive *–s*. However, there was an increase in production of copula *is* after therapy for the auxiliary. Similarly, Subject 2 showed no change from baseline until after the initiation of therapy for copula *is,* copula *was,* and possessive *–s*, but there was an increase in the production of auxiliary *is* and *was* after therapy for the comparable form of the copula. Hegde (1980) also found a significant crossover in production accuracy between copula and auxiliary. The reason for this crossover effect is unclear. Hegde concluded that auxiliary *BE* and copula *BE* represented a single response class to the children given the similarity in their surface form. If so, then the crossover could reflect a desirable generalization to an untrained exemplar.

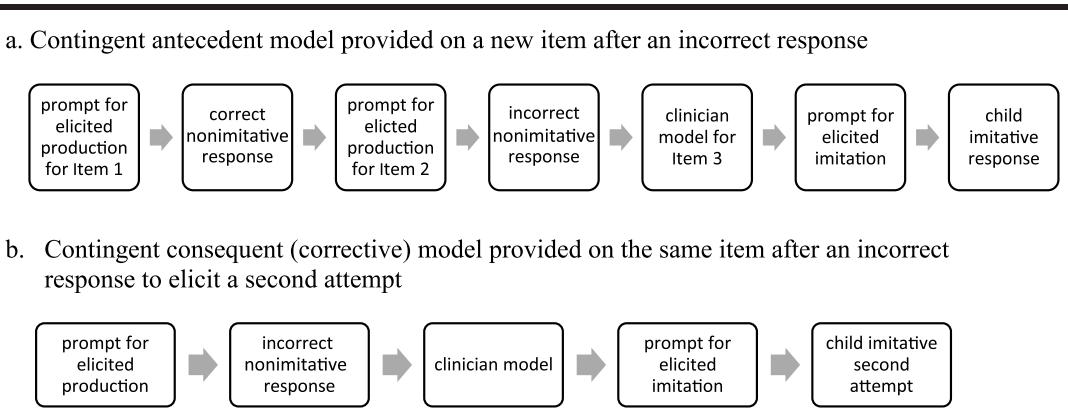
Figure 5. Timing of the elicitation prompt relative to the child's response.



A frequently mentioned limitation of imitation training is a lack of generalization to untrained exemplars or conversational speech (Fey, 1986; Paul & Norbury, 2012). Two studies (Hegde & Gierut, 1979; Zwitman & Sonderman, 1979) reported generalization to untrained exemplars (i.e., use of the targeted forms in different linguistic contexts). Hegde (1980) reported high levels of usage in conversational speech for copula and auxiliary once the child achieved 100% usage on therapy trials: ≥ 90% usage in conversation with the clinician in the clinic and 89% usage with the parent at home. However, Mulac and Tomlinson (1977) found no generalization to conversation without adding a procedure for transfer training with the parent at home.

Another type of generalization is to untrained members of the target category. In the study by Valian and Casey (2003), there was no generalization of question inversion to untrained auxiliaries when participants heard a single antecedent model before the elicited imitation attempt (per Figure 7a). However, generalization to untrained auxiliaries did occur when modeling density was increased so that participants either heard two antecedent models (per Figure 7b) or heard a model both before and after the elicited imitations (per Figure 7c).

Figure 6. Contingency of models.



Overall, the studies showed an increase in performance following imitation training therapy and higher performance relative to a control group that did not receive any therapy. It is important to note that, in all but one study, targets were absent (i.e., produced with 0% accuracy) prior to therapy. However, the studies do not unequivocally link the increased performance to the imitation training therapy or consistently show generalization of production to conversational speech or untrained exemplars.

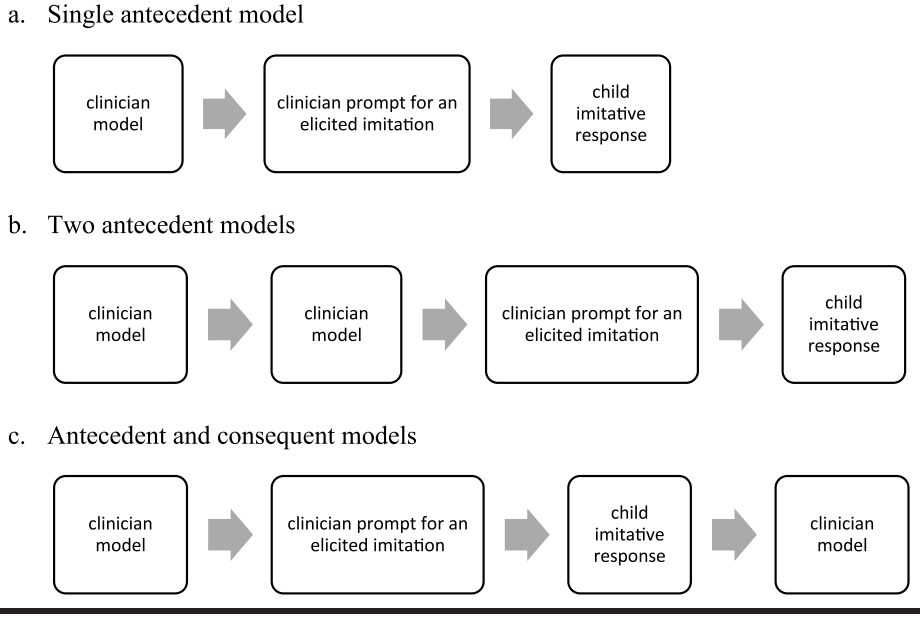
Studies Comparing Imitation Training to Observational Modeling

Imitation Training With Antecedent Models

Eight studies compared imitation training to observational modeling. Recall that, in imitation training, children both hear a model and attempt to produce the modeled utterance after the model, whereas in observational modeling, children hear a set of utterances with the target form (see Figure 3).

Three studies (Anderson, 2001; Connell, 1987; Connell & Stone, 1992) compared imitation training to observational modeling alone without any production attempts (illustrated in Figure 3b) for targets that were absent at baseline. In one of the studies, the child was instructed to just listen without saying anything during the observational modeling condition (Connell, 1987). In two other studies, the child produced a similar utterance that did not include the target form after hearing each model rather than just listening without saying anything (Anderson, 2001; Connell & Stone, 1992). To illustrate, in Connell and Stone (1992), the child first heard a model of the target form ("You tell

Figure 7. Variations in density of models (based on Valian & Casey, 2003).



Spinky to pick up the TVum.", where "um" was an invented morpheme that coded brokenness) and was then instructed to give a command to the puppet Spinky. In the imitation condition, the child was told "Say 'Spinky, pick up the TVum'" to elicit an imitation of the target form, whereas in the observational modeling condition, the child was told "Say 'Spinky, pick it up'" so that the child would not say the target form.

In both studies by Connell (1987; Connell & Stone, 1992), the rate of correct responding by children with DLD was higher in the imitation training condition than in the observational modeling condition, while the children with TL either had a higher accuracy rate for observational modeling (Connell, 1987) or showed no difference between conditions (Connell & Stone, 1992). In contrast, Anderson (2001) reported no difference in performance for children with DLD between imitation training and observational modeling.

In three other studies (Courtright & Courtright, 1976, 1979; Haynes & Haynes, 1980), observational modeling was followed with elicited production (see Figure 3c). The children were prompted to attempt production of the target—either the same utterances that had been modeled (Courtright & Courtright, 1976) or novel utterances (Courtright & Courtright, 1979; Haynes & Haynes, 1980) after hearing the set of models. Courtright and Courtright (1976) reported an initially higher rate of correct responding for imitation training followed by a plateau in performance. The observational modeling condition plus elicited production showed a lower initial rate of correct responding followed by a steady increase in accuracy and a higher final accuracy rate 1 week after the completion of the therapy sessions. It is important to note that baseline accuracy was higher for the imitation training group, ranging from 10% to 65% compared to 0% usage for the observational modeling group. Courtright and Courtright (1979) similarly reported a higher accuracy rate for the observational modeling plus elicited production condition at the 1-week follow-up but also found higher accuracy in each session for targets that were absent at baseline for both therapy conditions. In contrast, Haynes and Haynes (1980) reported a higher rate of correct responding in all sessions for imitation training but reported no difference at a 1-week follow-up for targets that were absent or emerging (i.e., produced with 10% accuracy) at baseline.

Overall, the studies contrasting imitation training with observational modeling alone (i.e., which did not include production of the target form) suggest that providing opportunities for children with DLD to attempt production of the target form will result in better performance. However, the studies do not unequivocally show imitation training to be more effective than observational modeling that includes production attempts. While some studies suggested that imitation training (i.e., having imitative attempts alternate with the model) might yield an initially higher rate of correct productions for modeled utterances, other studies showed that observational modeling with elicited production (i.e., having the child first hear a set of models before attempting nonimitative production of the target) may be equally effective in increasing accuracy for modeled utterances and might even result in higher accuracy over time for nonimitative production of untrained exemplars. Note that, in four of these studies (Anderson, 2001; Connell, 1987; Connell & Stone, 1992; Courtright & Courtright, 1976), the target was an invented grammatical form and we do not know whether this affected the outcome.

Imitation Training With Consequent Models

Only one study (Goldstein, 1984) compared imitation training with consequent corrective models to observational modeling without production attempts. This study included only children with TL and targeted an invented grammatical form. In Goldstein (1984), two children only learned the target forms in the imitation training condition. The other four children showed more rapid acquisition of the target forms in the imitation training condition than in the observational modeling condition. This was consistent with the studies of antecedent models; that is, performance was better when children had to attempt production of the target form. There were no studies that compared imitation training with consequent models to modeling followed by production attempts.

Studies Comparing Imitation Training to Embedded Therapy

Imitation training provides concentrated opportunities for the child to hear models of a target form and attempt that form outside of a communicative interaction. Six studies (Camarata & Nelson, 1992; Camarata et al., 1994; Cole & Dale, 1986; Friedman & Friedman, 1980; Hughes & Carpenter, 1989; Nelson et al., 1996) compared imitation training to embedded therapy approaches such as interactive language development teaching (ILDT; Lee, 1974) and conversational recast therapy (CRT; Baker & Nelson, 1984). Within these embedded therapy approaches, the child hears and attempts the target form for communicative purposes. However, the density of models and attempts is lower than in imitation training.

ILDT is implemented during storytelling activities and includes elicited imitation as one component of therapy. Models for imitation are provided both prior to child attempts and after child errors. Hughes and Carpenter (1989) reported a slight advantage for imitation training over ILDT—performance for three out of four participants was higher in the imitation training condition. All four of the targets in the imitation training condition reached criterion compared to three of the targets in ILDT. However, these authors did not evaluate whether the outcomes were statistically different. In addition, only mean performance was reported, and there was no indication about whether outcome varied as a function of baseline levels, which varied from 0% to 26% usage. There was little generalization to conversation for either therapy.

In other studies, conversational recasting was implemented during play activities. Children heard models only after inaccurate attempts at the target form. These consequent

corrective models (termed "recasts") added the missing target form to the child's utterance, and there was no prompting for the child to imitate those models. In these studies, authors looked at both targeted productions during therapy activities (defined as imitative, elicited, or spontaneous productions of therapy targets during the therapy session in which that form was targeted) and nontargeted productions (defined as elicited or spontaneous nonimitative productions of the target form during a session in which that form was not being targeted).[2] For targeted production during therapy sessions, there was either no difference between therapy conditions (Camarata & Nelson, 1992) or an advantage for imitation training (Camarata et al., 1994; Nelson et al., 1996). In these latter studies, children required fewer sessions and fewer therapy trials in the imitation training condition before producing the first targeted production of a grammatical form. In both Camarata and Nelson (1992) and Camarata et al. (1994), the first nontargeted production was achieved after fewer sessions and with fewer therapy trials for recasting than for imitation training. Furthermore, the time between achieving the first targeted production and the first nontargeted production was much shorter for conversational recasting than for imitation training (Camarata & Nelson, 1992), and some children did not achieve nontargeted production during the imitation training condition (Nelson et al., 1996). In contrast, Gillum et al. (2003) reported an advantage for conversational recasting for achieving a higher criterion on targeted productions (defined as five productions) as well as on achieving criterion for nontargeted production. Note that all of these studies included targets that were absent at baseline. Nelson et al. (1996) also included partially mastered targets, and these targets showed a larger increase in nontargeted production for the recasting condition.

Two studies looked at whether child differences in cognitive level (as measured by IQ) or language predicted which therapy would be more effective. The outcome measure for both studies was the Developmental Sentence Score (DSS; Lee, 1974) based on a conversational sample, and neither study specified the baseline level for therapy targets. In a study by Friedman and Friedman (1980) comparing imitation training to ILDT, children showed moderate gains in DSS after both therapy approaches, regardless of pretherapy language level. However, there was an interaction of therapy condition with IQ. Children with lower IQ showed higher gains for imitation training, while children with higher IQ made more gains in the ILDT condition. In contrast, Cole and Dale (1986) found no difference in DSS gain between imitation training and a conversationally based therapy approach that was similar to CRT, and neither IQ nor pretherapy language level predicted the amount of gain from either therapy.

[2]In their reports, these authors used the terms "elicited production" and "spontaneous production." However, to avoid confusion and maintain consistency with our use of terminology, we instead used the terms "targeted production" and "nontargeted production."

Overall, the studies suggest that imitation training might be less effective for long-term learning than a conversationally based therapy approach for children with DLD. Children might start to produce the target form more quickly during imitation training. However, achieving a higher level of production in therapy and generalizing usage to conversational speech might be achieved more quickly with a conversationally based therapy approach. It is also, however, worth noting the short-term effect of imitation training on performance. Imitation training led to a rapid increase in production during therapy for target forms that were completely absent prior to therapy. It is also important to note that the advantage of conversationally based therapy might not hold for all grammatical forms. Camarata and Nelson (1992), for instance, noted that nontargeted production of gerunds was achieved more quickly in the imitation training condition.

Discussion

The reviewed studies show that imitation training increased production of grammatical forms in therapy. Specifically, children quickly learned to repeat the modeled utterances during therapy trials and generalized production to untrained exemplars. Importantly, this effect was seen for grammatical forms that were absent prior to therapy. The strongest evidence for this effect was from two Level 2 group studies that compared imitation training to CRT (Camarata et al., 1994; Nelson et al., 1996).

Imitation training involves two major components—hearing models of the target form and production attempts. The review suggests that production attempts are an important component of therapy (Anderson, 2001; Connell, 1987; Connell & Stone, 1992; Goldstein, 1984). However, it also suggests that production attempts need not be imitative (Camarata & Nelson, 1992; Camarata et al., 1994; Courtright & Courtright, 1976, 1979; Nelson et al., 1996).

We considered whether there was any best way to implement imitation training. We found only one study that compared different ways of implementing imitation training. This study showed increased efficacy for imitation training when the child heard more models (Valian & Casey, 2003). This apparent benefit of increasing the density of models might explain why observational modeling was at least as effective as if not more so than, imitation training when observational modeling included production attempts (Courtright & Courtright, 1976, 1979). However, the timing of those models might not be important. Specifically, it might not be necessary to provide models before production attempts, as was done in most of the studies. Consequent models have also been shown to be effective in increasing production of target forms (Goldstein, 1984).

Contingency of models might also not be crucial. Performance increased when antecedent models were provided before every attempt, regardless of accuracy, as was done in many of the studies. However, starting antecedent models only after errors and then eliminating the model after some number of correct was also effective (Hegde,

1980; Hegde & Gierut, 1979; Hegde et al., 1979; Hughes & Carpenter, 1989). Providing consequent models only after errors was also effective (Goldstein, 1984). The review showed that children may make fewer errors when models are given before every attempt. Yet, the success of contingent models suggests that errorless learning is not a necessary condition for children to improve in therapy.

There has been a long-standing concern about a potential lack of generalization to conversational speech for imitation training (e.g., Fey, 1986; Leonard, 1981; Paul & Norbury, 2012). Many of the studies did not evaluate this. Those that did confirm that this may be a limitation. Generalization to conversational speech may not occur without the implementation of procedures to achieve the transfer of usage to conversational speech contexts (Mulac & Tomlinson, 1977) or after considerably higher amounts of therapy than therapies that involved conversational contexts (Camarata & Nelson, 1992; Camarata et al., 1994). Thus, the benefit of imitation training—a rapid increase in the short term for usage in therapy—might be countered by a slower rate of change for achieving long-term goals for communicative use.

An important consideration is to match therapy approaches to the client as therapy approaches might not be equally effective for all children. Our review suggests that cognitive level might be important to consider. That is, imitation training might be particularly useful for children with lower cognitive levels (Friedman & Friedman, 1980). Our review also suggests that how children with TL respond to therapy might not be applicable to children with DLD. That is, imitation training may be useful to help children with DLD rapidly achieve production of target forms even if children with TL achieve production more readily with another therapy approach (Connell, 1987; Connell & Stone, 1992).

Therapy approaches might also not be equally effective for all grammatical forms. Of note is the finding by Camarata and Nelson (1992) that nontargeted production of gerunds, a low-frequency occurring form, was achieved faster with imitation training than with recast therapy. Imitation training might be particularly useful for such infrequently occurring forms. This may explain the higher accuracy rates for imitation training than for observational modeling in studies of invented morphemes (Connell, 1987; Connell & Stone, 1992).

Despite a long history of using imitation training for working on grammar goals, there has been little investigation into the best way for implementing this therapy approach to maximize efficacy. There are alternative ways of implementing imitation training, and we need studies to determine which implementation conditions might result not only in increased usage in therapy but also in increased generalization to untrained exemplars and to conversational usage. Several lines of inquiry would be valuable, including comparing partial to exact imitation, comparing immediate to delayed imitation, comparing antecedent to consequent models, and comparing contingent to noncontingent models (i.e., models given regardless of accuracy). We also need studies that manipulate the density of models and explore how this affects the relative efficacy of imitative versus nonimitative production responses.

There appears to be little recent interest in investigating the efficacy of imitation training. This may reflect an assumption that the efficacy of this therapy approach had already been established and that additional studies are not needed. Our review shows that this is not the case. The studies whose aim was to evaluate imitation training, all of which were from 1980 or earlier, did not unequivocally prove the efficacy of imitation training. The focus of more recent studies has been to show that other therapy approaches work better than imitation training. Interestingly, it is these studies—particularly the comparisons with CRT—that have been most illuminating about the impact of imitation on the production of grammatical forms.

In phonological disorders, researchers have found that increasing a child's ability to imitate target sounds (termed "stimulability") leads to greater gains in therapy (Miccio & Elbert, 1996; Rvachew et al., 1999). The same may be true for grammatical targets. That is, the rapid achievement of imitative production from imitation training might enhance the outcome of the other therapy approaches and lead to greater gains, particularly for forms that are absent at the start of therapy.

Therapy studies are typically set up in an either-or scenario to determine whether one therapy works better than another. Studies are also needed to investigate whether combining two therapy approaches would work better than using only one of the approaches. For instance, given the difference in immediate and long-term outcomes between imitation training and CRT, it may be that a therapy protocol that combines these two approaches would be more effective than either approach alone.

Table 10. Imitation training variables providing more and less support.

Variables	More support	Less support
Degree of match	Exact imitation	Partial imitation
Timing of prompt	Immediate imitation	Delayed imitation
Timing of model	Antecedent model	Consequent model
Density of models	Before every attempt	Reduced schedule of models (i.e., before every second, third, … attempt)
Contingency of model	Provide model regardless of correctness	Provide model only after error

Concluding Thoughts

Applying evidence to clinical practice is a challenge. To be confident in the value of using a particular therapy approach, we need evidence that the therapy approach works. We have some evidence for this for imitation training, at least regarding production of target forms during therapy activities and for extension of production beyond the trained exemplars. We also want evidence that the therapy approach works better than other therapies and that, therefore, it is the best choice of approaches for working on a particular target with a particular client. The evidence regarding this is mixed for imitation training. Imitation training may be the fastest way to achieve initial production and a high rate of accurate production during imitative therapy trials. However, it may not be the most efficient approach for achieving the longer term aim of communicative use in conversation.

This does not mean that clinicians need to completely stop using imitation training. Rather, it suggests the need to modify how imitation training is implemented. Specifically, we may want to discontinue the practice of first training a form in imitation until a high criterion is reached before working on that same form within discourse. An alternative implementation would combine imitation training with conversationally based therapies, by first practicing the target during a brief imitation training activity and then immediately following that with an activity involving communicative use of the target form (Eisenberg, 2005, 2013, 2014). Such a combined approach has been successfully used in several studies (Fey et al., 1993; Owen Van Horne et al., 2018). This approach might be particularly useful for grammatical forms that occur less frequently in conversation.

Clinicians might also want to reconceptualize imitation training as involving a continuum of support rather than as an all-or-none phenomenon. In most studies, models for immediate exact imitation were provided before every attempt and then abruptly discontinued when the child achieved some criterion for imitative production on therapy trials. It may be that the lack of generalization may be due to this abrupt elimination of the model rather than to imitation training per se. An alternative would be to manipulate variables such as the degree of match between child response and model, timing of prompts and models, contingency of models, and density of models to reduce support as shown in Table 10. This would provide a hierarchy of scaffolding that could be individualized to client needs.

In conclusion, we believe that the use of imitation training, with suggested modifications, can continue to be a useful tool in an SLP's toolbox. Further studies, such as those suggested above, will serve to refine the procedure and add to the evidence base to support our expert opinion.

References

Anderson, R. T. (2001). Learning an invented inflectional morpheme in Spanish by children with typical language skills and with specific language impairment (SLI). *International Journal of Language & Communication Disorders, 36*(1), 1–19.

Baker, N., & Nelson, K. E. (1984). Recasting and related conversational techniques for triggering syntactic advances by young children. *First Language, 5*(13), 3–21.

Bates, E., Benigni, L., Bretherton, I., Camioni, L., & Volterra, V. (1979). *The emergence of symbols: Cognition and communication in infancy*. Academic Press.

Bates, E., Thal, D., Whitesell, K., Fenson, L., & Oakes, L. (1989). Integrating language and gesture in infancy. *Developmental Psychology, 25*(6), 1004–1019.

Bloom, L., Hood, L., & Lightbown, P. (1974). Imitation in language development: If, when, and why. *Cognitive Psychology, 6*(3), 380–420.

Bredin-Oja, S. L., & Fey, M. E. (2014). Children's responses to telegraphic and grammatically complete prompts to imitate. *American Journal of Speech-Language Pathology, 23*(1), 15–26. https://doi.org/10.1044/1058-0360(2013/12-0155)

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*(3), 167–178.

Camarata, S. M., Nelson, K. E., & Camarata, M. N. (1994). Comparison of conversational-recasting procedures and imitative procedures for training grammatical structures in children with specific language impairment. *Journal of Speech and Hearing Research, 37*(6), 1414–1423. https://doi.org/10.1044/jshr.3706.1414

Cole, K. N., & Dale, P. S. (1986). Direct language instruction and interactive language instruction with language delayed preschool children: A comparison study. *Journal of Speech and Hearing Research, 29*(2), 206–217. https://doi.org/10.1044/jshr.2902.206

Connell, P. J. (1987). An effect of modeling and imitation teaching procedures on children with and without specific language impairment. *Journal of Speech and Hearing Research, 30*(1), 105–113. https://doi.org/10.1044/jshr.3001.105

Connell, P. J., & Stone, C. A. (1992). Morpheme learning of children with specific language impairment under controlled instructional conditions. *Journal of Speech and Hearing Research, 35*(4), 844–852. https://doi.org/10.1044/jshr.3504.844

Courtright, J. A., & Courtright, I. C. (1976). Imitative modeling as a theoretical base for instructing language-disordered children. *Journal of Speech and Hearing Research, 19*(4), 655–663. https://doi.org/10.1044/jshr.1904.655

Courtright, J. A., & Courtright, I. C. (1979). Imitative modeling as a language intervention strategy: The effects of two mediating variables. *Journal of Speech and Hearing Research, 22*(2), 389–402. https://doi.org/10.1044/jshr.2202.389

Darrah, J., Hickman, R., O'Donnell, M., Vogtle, L., & Wiart, L. (2008). *AACPDM methodology to develop systematic reviews of treatment interventions (Revision 1.2)*. http://www.aacpdm.org/UserFiles/file/systematic-review-methodology.pdf

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*(1), 7–40.

Eisenberg, S. (2005). When conversation is not enough: Assessing infinitival complements through elicitation. *American Journal of Speech-Language Pathology, 14*(2), 92–106. https://doi.org/10.1044/1058-0360(2005/011)

Eisenberg, S. (2013). Grammar intervention: Content and procedures for facilitating children's language development. *Topics in Language Disorders, 33*(2), 165–178.

Eisenberg, S. (2014). What works in therapy: Further thoughts on improving clinical practice for children with language disorders. *Language, Speech, and Hearing Services in Schools, 45*(2), 117–126. https://doi.org/10.1044/2014_LSHSS-14-0021

Fey, M. E. (1986). *Language intervention with young children*. College-Hill Press.

Fey, M. E., Cleave, P. L., Long, S. H., & Hughes, D. L. (1993). Two approaches to the facilitation of grammar in children with language impairment: An experimental evaluation. *Journal of Speech and Hearing Research, 36*(1), 141–157. https://doi.org/10.1044/jshr.3601.141

Fey, M. E., & Proctor-Williams, K. (2000). Recasting, elicited imitation, and modelling in grammar intervention for children with specific language impairments. In D. V. H. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention, and outcome* (pp. 177–194). Psychology Press.

Finestack, L. H., & Satterlund, K. E. (2018). Current practice of child grammar intervention: A survey of speech-language pathologists. *American Journal of Speech-Language Pathology, 27*(4), 1329–1351. https://doi.org/10.1044/2018_AJSLP-17-0168

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior, 2*(2), 121–135.

Friedman, P., & Friedman, K. A. (1980). Accounting for individual differences when comparing the effectiveness of remedial language teaching methods. *Applied Psycholinguistics, 1*(2), 151–170.

Gillum, H., Camarata, S., Nelson, K. E., & Camarata, M. N. (2003). A comparison of naturalistic and analog treatment effects in children with expressive language disorder and poor preintervention imitation skills. *Journal of Positive Behavior Interventions, 5*(3), 171–178.

Goldstein, H. (1984). Effects of modeling and corrected practice in generative language learning of preschool children. *Journal of Speech and Hearing Disorders, 49*(4), 389–398. https://doi.org/10.1044/jshd.4904.389

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews. *European Journal of Clinical Investigations, 48*(6), e12931.

Haynes, W. O., & Haynes, M. D. (1980). A comparison of nonimitative modeling and mimicry procedures for training the singular copula in Black preschool children. *Journal of Communication Disorders, 13*(4), 277–288.

Hegde, M. N. (1980). An experimental-clinical analysis of grammatical and behavioral distinctions between verbal auxiliary and copula. *Journal of Speech and Hearing Research, 23*(4), 864–877. https://doi.org/10.1044/jshr.2304.864

Hegde, M. N., & Gierut, J. (1979). The operant training and generalization of pronouns and a verb form in a languagedelayed child. *Journal of Communication Disorders, 12*(1), 23–34.

Hegde, M. N., Noll, M. J., & Pecora, R. (1979). A study of some factors affecting generalization of language training. *Journal of Speech and Hearing Disorders, 44*(3), 301–320. https://doi.org/10.1044/jshd.4403.301

Hughes, D. L., & Carpenter, R. L. (1989). Short-term syntax learning and generalization effects. *Child Language Teaching and Therapy, 5*(2), 202–219.

Lee, L. L. (1974). *Developmental sentence analysis*. Northwestern University Press.

Leonard, L. B. (1981). Facilitating linguistic skills in children with specific language impairment. *Applied Psycholinguistics, 2*(2), 89–118.

Leonard, L. B. (2014). *Children with specific language impairment*. MIT Press.

MacLure, M. (2005). Clarity bordering on stupidity: Where's the quality in systematic review? *Journal of Education Policy, 20*(4), 393–416.

Matheny, N., & Panagos, J. M. (1978). Comparing the effects of articulation and syntax programs on syntax and articulation improvement. *Language, Speech, and Hearing Services in Schools, 9*(1), 57–61. https://doi.org/10.1044/0161-1461.0901.57

Meltzoff, A. N., & Williamson, R. A. (2013). Imitation: Social, cognitive, and theoretical perspectives. In P. D. Zelazo (Ed.), *Oxford handbook of developmental psychology* (Vol. 1, pp. 651–682). Oxford University Press.

Miccio, A. W., & Elbert, M. (1996). Enhancing stimulability: A treatment program. *Journal of Communication Disorders, 29*(4), 335–351.

Mulac, A., & Tomlinson, C. N. (1977). Generalization of an operant remediation program for syntax with language delayed children. *Journal of Communication Disorders, 10*(3), 231–243.

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*(4), 850–859. https://doi.org/10.1044/jshr.3904.850

Owen Van Horne, A. J., Curran, M., Larson, C., & Fey, M. E. (2018). Effects of a complexity-based approach on generalization of past tense *–ed* and related morphemes. *Language, Speech, and Hearing Services in Schools, 49*(Suppl. 3), 681–693.

Paul, R., & Norbury, C. F. (2012). *Language disorders from infancy through adolescence: Listening, speaking, reading, writing, and communicating*. Elsevier.

Rvachew, S., Rafaat, S., & Martin, M. (1999). Stimulability, speech perception skills, and the treatment of phonological disorders. *American Journal of Speech-Language Pathology, 8*(1), 33–43. https://doi.org/10.1044/1058-0360.0801.33

Speidel, G. E., & Herreshoff, M. J. (1989). Imitation and the construction of long utterances. In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 181–197). Springer-Verlag.

Tager-Flusberg, H., & Calkins, S. (1990). Does imitation facilitate the acquisition of grammar? Evidence from a study of autistic, Down syndrome and normal children. *Journal of Child Language, 17*(3), 591–606.

Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.

Valian, V., & Casey, L. (2003). Young children's acquisition of wh-questions: The role of structured input. *Journal of Child Language, 30*(1), 117–143.

Zwitman, D. H., & Sonderman, J. C. (1979). A syntax program designed to present base linguistic structures to language-disordered children. *Journal of Communication Disorders, 12*(4), 323–335.