Effects of Augmentative and Alternative Communication Intervention on Speech Production in Children With Autism: A Systematic Review

Ralf W. Schlosser
Northeastern University, Boston

Oliver Wendt
Purdue University, West Lafayette, IN

Purpose: This systematic review aimed to determine the effects of augmentative and alternative communication (AAC) intervention on speech production in children with autism or pervasive developmental disorder-not otherwise specified.
Method: A systematic review methodology was utilized to limit bias in searching, selecting, coding, and synthesizing relevant treatment studies. This involved a multifaceted search for studies written between 1975 and May 2007 using various bibliographic databases, dissertation databases, hand searches of selected journals and published compilations of AAC theses and dissertations, and ancestry searches. To be included, studies had to meet stringent criteria. A coding manual and form facilitated data extraction in terms of participant characteristics, treatment characteristics, design and measurement, and outcomes.
Results: Nine single-subject experimental design (27 participants) and 2 group studies (98 participants) were included. Results indicated that AAC interventions do not impede speech production. In fact, most studies reported an increase in speech production. However, in-depth analyses revealed that the gains were rather modest.
Conclusions: Although AAC interventions do not appear to impede speech production and may result in increased speech production, the modest gains observed require realistic expectations among clinicians and other stakeholders. Future research should be more hypothesis driven and aim to identify predictive child characteristics, such as prior speech imitation and object exploration skills.

Key Words: augmentative and alternative communication, autism spectrum disorders, speech improvement, systematic review

Approximately 25%–61% of learners with autism present with little or no functional speech (Weitz, Dexter, & Moore, 1997) and may be candidates for augmentative and alternative communication (AAC) approaches to replace or supplement natural speech and or handwriting (Lloyd, Fuller, & Arvidson, 1997). Unaided AAC approaches include gestures, manual signing, and finger spelling. Aided AAC approaches include selection-based methods, such as graphic symbols, nonelectronic communication boards, speech-generating devices with synthesized and/or digitized speech output, and exchange-based approaches, such as the Picture Exchange Communication System (e.g., Mirenda, 2003). Although there is no shortage of potential AAC approaches for individuals with autism, many families of children with autism as well as speech-language pathologists and other service providers are concerned that the use of AAC may hinder their child's natural speech production (Schlosser, 2003). On the other hand, several authors have argued that AAC interventions might actually facilitate speech production (Blischak, Lombardino, & Dyson, 2003; Frost & Bondy, 2002; M. Sundberg, Michael, Partington, & Sundberg, 1996).

It is understood that the primary aim of AAC intervention is to facilitate a child's communicative competence through the use of multiple communication modalities that are by their very nature supplementing ("augmentative") or replacing ("alternative") natural speech (Light, Beukelman, & Reichle, 2003). Thus, although improvements in speech production

1058-0360/08/1703-0212

per se are not a primary goal of AAC interventions, such outcomes do represent a welcomed bonus to AAC intervention efforts. Knowing whether AAC interventions facilitate or hinder speech production has the potential to inform current AAC intervention practices for children with autism in several ways. First, should AAC interventions be found not to hinder speech production (or even to enhance speech production), practitioners and families might be more inclined to support earlier adoption of AAC modalities and interventions rather than rely on a "wait-and-see" approach for children with limited or no functional speech. Second, if data were to show that some AAC interventions yield better speech production outcomes than others, practitioners and families could consider this evidence in their treatment selection decisions.

Thankfully, in several AAC intervention studies, the investigators have not only monitored the impact on communication skills but also evaluated the effects of AAC interventions on speech production. Any one individual study, however, cannot unequivocally reveal whether AAC intervention does facilitate speech production. Individual studies may produce certain results simply by chance, suggesting the need for a synthesis of studies. Several authors have aimed to synthesize the evidence on AAC interventions for individuals with autism (Goldstein, 2002; Koul, Schlosser, & Sanscribrian, 2001; Mirenda, 2001, 2003; Mirenda & Eriksson, 2000; Paul, 2003; Potter & Brown, 1997; Schlosser & Blischak, 2001; Sigafoos & Drasgow, 2001; C. Sundberg & Michael, 2001; M. Sundberg, 1993). Although these reviews offered a valuable service by synthesizing information across studies, none of them focused solely on the effects of AAC intervention on natural speech production. Moreover, the authors of these reviews provided narrative descriptions of the evidence. Narrative reviews have a tendency to be subjective and are prone to various biases, such as search and selection bias. Hence, they differ considerably from systematic reviews (Schlosser & Goetze, 1992):

> Systematic reviews are literature reviews that adhere closely to a set of scientific methods that explicitly aim to limit systematic error (bias), mainly by attempting to identify, appraise and synthesize all relevant studies (of whatever design) in order to answer a particular question (or set of questions). (Petticrew & Roberts, 2006, p. 9)

A computerized database search revealed three systematic reviews that bear relevance to the topic at hand (Millar, Light, & Schlosser, 2006; J. B. Schwartz & Nye, 2006; Wendt, 2007). J. B. Schwartz and Nye (2006) evaluated the effects of manual sign intervention (sign alone or total communication) on signed and spoken communication outcomes in children with autism between 4 and 18 years of age. Although systematic in approach, this review is limited for our purposes because the included studies reported mostly signed and spoken (together) outcomes rather than spoken communication per se. Thus, the aggregated data are not relevant to our question. Only two of the eight included studies bear relevance (M. Cohen, 1979; Yoder & Layton, 1988). Moreover, this review included only manual signing, which is just one of many AAC approaches as described earlier.

Wendt (2007) aimed to synthesize all AAC interventions for children with autism spectrum disorders in terms of a number of outcomes. One of these outcomes was speech production, for which he identified seven studies. However, neither the inclusion criteria nor the coding categories were sufficiently specific to speech production because this represented only one of many dependent variables. Nonetheless, there was some overlap with and some differences from the current effort in terms of the included studies.

Finally, Millar et al. (2006) focused directly on the effects of AAC interventions on speech production in children with developmental disabilities. This included children with autism but also children with intellectual disabilities, cerebral palsy, and so forth. Six studies, involving 27 participants, met their criteria for inclusion for a best evidence analysis. None of the 27 participants demonstrated decreases in speech production, and 89% showed gains in speech production. Because the population targeted was not restricted to autism (in fact, only two studies of the best evidence analysis involved children with autism), the coded participant characteristics were kept more general. Additionally, the results were not analyzed separately for children with autism, and thus it is unclear whether the overall conclusions were applicable to children with autism per se. Also, only studies up to 2003 were included, and so more recent evidence did not enter this review. Finally, Millar et al. (2006) required that the study had to provide data not only on speech production but also on the acquisition of an AAC modality. This may have resulted in the exclusion of some studies that provided only speech production data. In order to determine whether AAC intervention rather than AAC acquisition results in increased speech production, AAC acquisition data are not necessary. Thus, there are sufficient reasons for a focused systematic review on the effects of AAC interventions on speech productions in children with autism or pervasive developmental disorder-not otherwise specified (PDD-NOS).

Method

Criteria for Inclusion

In order to be included in this review, a study needed to meet the following criteria: (a) the intervention had to be classified as within the scope of AAC, defined as the following (American Speech-Language-Hearing Association, 2002, p. 98):

> a set of procedures and processes by which an individual's communication skills (i.e., production as well as comprehension) can be maximized for functional and effective communication. It involves supplementing or replacing natural speech and/or writing with aided (e.g., picture communication symbols, line drawings, Blissymbols, and tangible objects) and/or unaided symbols (e.g., manual signs, gestures, and finger spelling).

(b) the intervention did not involve functional communication training (FCT) with AAC because the primary goal of these interventions is to replace challenging behaviors with appropriate communicative forms that serve the same function as the challenging behavior, and hence, FCT studies involving children with little or no functional speech typically monitor challenging behaviors and appropriate communicative behaviors but not speech production as a secondary outcome to the use of AAC (see Bopp, Brown, & Mirenda, 2004); (c) the participants were not functionally speaking prior to AAC intervention; (d) speech production was monitored as a dependent variable (anecdotal reports of changes in speech production did not qualify; similarly, studies that included speech as only one of several modalities constituting a correct response were excluded; studies in which the dependent measure requires natural speech and an AAC mode or natural speech and/or an AAC mode in order to be counted as correct were excluded as well [e.g., Layton, 1988]); (e) data for speech production from single-subject experimental designs needed to lend themselves for the calculation of the percentage of nonoverlapping data (PND; i.e., graphic display of session-by-session time series data, no ceiling effects in the baseline), and data from group designs needed to permit the calculation of effect size (ES); (f) participants had a diagnosis of autism or PDD-NOS (participants with concomitant disabilities such as intellectual disabilities qualified as well as long as they also had a diagnosis of autism or PDD-NOS; participants described as "autistic-like" [e.g., Sigafoos & Drasgow, 2001] or participants with other disabilities without autism or PDD-NOS did not qualify [e.g., Stoner et al., 2006]); (g) the target of the intervention had to be the child with autism or PDD-NOS rather than communication partners even when the child's speech production was monitored as a dependent measure (e.g., Howlin, Gordon, Pasco, Wade, & Charman, 2007); (h) quasi-experimental designs, such as selected group designs or single-subject experimental designs, were used for evaluating an intervention (pre-experimental designs, such as AB designs, or group equivalents, such as pretest–posttest designs, were excluded [e.g., I. S. Schwartz, Garfinkle, & Bauer, 1998]); (i) in a group design, all subjects had to be classified as autistic or PDD-NOS, and more than 90% of the enrolled subjects had to enter statistical analyses (if some subjects had other developmental disabilities and there were no separate analyses for subjects with autism, the study did not qualify [e.g., DiCarlo, Stricklin, Banajee, & Reid, 2002]; similarly, if 90% or fewer subjects entered statistical analysis, the study was excluded [Carr & Felce, 2007a]); (j) the study was published in a peer-reviewed journal or approved as a dissertation or thesis (conference papers from proceedings [e.g., M. Cohen, 1979] did not qualify because they typically do not offer enough detail for the coding within a systematic review); and (k) the study had to be dated between 1975 and May 2007 (the cutoff year 1975 was selected because this was approximately the time when manual signing started to be used with children with autism [e.g., Carr, Binkoff, Kologinsky, & Eddy, 1978], and it marks the beginning of the use of single-subject experimental designs as a research methodology for evaluating treatment effectiveness [Barlow & Hersen, 1973]).

Locating Studies

A multifaceted search strategy was implemented following published tutorials on searching for evidence (Schlosser, Wendt, Angermeier, & Shetty, 2005; Schlosser, Wendt, Bhavnani, & Nail-Chiwetalu, 2006; White, 1994). First, the reference lists of previously cited published and unpublished reviews related to AAC and autism were examined. This yielded several potential studies (e.g., Charlop-Christy, Carpenter, Le, LeBlanc, & Kellet, 2002). Second, several general-purpose databases were consulted, including the Cumulative Index of Nursing and Allied Health (CINAHL), Educational Resources Information Clearinghouse (ERIC), Linguistics and Language Behavior Abstracts (LLBA), Medline, and PsycINFO. The search terms used were specific to each database. In order to arrive at workable key words, the authors relied on the Pearl Growing strategy (Schlosser et al., 2006): (a) find a relevant article (here: Charlop-Christy et al., 2002); (b) find the terms under which the article is indexed in Database 1; (c) find other relevant articles in Database 1 by using the index terms in a Building Block query; (d) repeat b and c in other databases; (e) repeat steps a–d for other relevant articles; and (f) end when articles retrieved provide diminishing relevance. On the basis of the terms under which Charlop-Christy et al. (2002) was indexed in PsycINFO, Query 1 in Table 1 was devised. Because this yielded only three references (including the pearl), we dropped the key word "speech development" as speech production is often a secondary consideration in AAC interventions, and therefore, a study may not be indexed as such. Instead, we replaced it with potential free-text words that might capture this dependent variable. Then we ran the search again, which increased the yield substantially. This was subsequently repeated in Medline (see Table 1). The pearl was not indexed in CINAHL, ERIC, or LLBA. However, similar searches were constructed for these databases.

Third, database searches were conducted to identify doctoral dissertations and master's theses specifically. With the exception of Wendt (2007), dissertations did not qualify for inclusion in any of the previously identified reviews. The above search in PsycINFO had already yielded a dissertation for potential inclusion (i.e., Anderson, 2002), but PsycINFO indexes only psychology-related dissertations. Therefore, a search of ProQuest Digital Dissertations yielded several additional references with relevance to the topic (e.g., Ferrara, 2005; Ganz, 2002; Son, 2006; Tincani, 2003). Three of these dissertations, however, have been subsequently published in the peer-reviewed literature (Ganz & Simpson, 2004; Son, Sigafoos, O'Reilly, & Lancioni, 2006; Tincani, 2004). Fourth, publisher-related databases (www.sciencedirect.com, www.springerlink.com, www.metapress.com) were searched for articles that were in press or appeared first electronically before they were published in print. This resulted in several additional potential studies and reviews, some of which are now published (Carr & Felce, 2007a, 2007b; Ganz, Simpson, & Corbin-Newsome, 2007; Lancioni et al., 2007; Olive et al., 2007). Finally, active researchers in AAC and autism were contacted via e-mail, presented with a list of heretofore included and excluded studies, and asked to nominate additional studies. This resulted in several additional nominations (Koita & Sonoyama, 2004; Koita, Sonoyama, & Takeuchi, 2003; Sigafoos & Drasgow, 2001; Yokoyama, Naoi, & Yakamoto, 2006).

TABLE 1. Search strategies across general purpose databases.

Databases	Search strategy	Yield	Revised search
PsycINFO	(DE "Augmentative Communication" OR DE "Communication Skills Training" OR DE "Communication Systems" OR DE "Sign Language") and (DE "Autism") AND (DE "speech development") AND "Empirical Study" as a limiter	3 ref	Replaced key word "speech development" with relevant free-text words
	(DE "Augmentative Communication" OR DE "Communication Skills Training" OR DE "Communication Systems" OR DE "Sign Language") and (DE "Autism") AND (TX "speech" OR TX "vocal*" OR TX "spoken") AND the limiter "Empirical Study"	96 ref	
Medline	(MH "Communication Aids for Disabled" OR MH "Sign Language") AND (MH "Speech") AND (MH "Autistic Disorder")	6 ref	Dropped "speech"
	(MH "Communication Aids for Disabled" OR MH "Sign Language") AND (MH "Autistic Disorder")	95 ref	
ERIC	(DE "Augmentative and Alternative Communication" OR DE "Sign Language") AND (DE "Autism") AND (DE "Speech")	4 ref	Dropped "speech"
	(DE "Augmentative and Alternative Communication" OR DE "Sign Language") AND (DE "Autism")	126 ref	Added free-text words
	(DE "Augmentative and Alternative Communication" OR DE "Sign Language") AND (DE "Autism") AND (TX "speech" OR TX "vocal*" OR TX "spoken")	40 ref	
CINAHL	(MH "Alternative and Augmentative Communication" OR MH "Sign Language") AND (MH "Autistic Disorder") AND (MH "Speech")	2 ref	Dropped "speech"
	(MH "Alternative and Augmentative Communication" OR MH "Sign Language") AND (MH "Autistic Disorder")	30 ref	
LLBA	("Augmentative and alternative communication") AND ("autism") AND ("speech")	26 ref	

Note. ref = references; DE = descriptor; TX = free text; MH = major heading; ERIC = Educational Resources Information Clearinghouse; CINAHL = Cumulative Index of Nursing and Allied Health; LLBA = Linguistics and Language Behavior Abstracts. Asterisk indicates truncation that yields words which contain "vocal" as a root but may have various endings such as "vocalize," "vocalization," and "vocalized."

Potential studies were evaluated against the inclusion checklist, which contained the items listed above. The first and second authors independently completed the inclusion checklist for each potential study. There were only two disagreements that related to an uncertainty about whether studies that did not separate teaching from testing should be excluded. Following discussion, we resolved this issue by deciding that this would not lead to exclusion but would be considered in the appraisal of the study. This resulted in an agreement of 100%. Studies that were included are provided in the reference list (identified through an asterisk). A log of excluded studies, which allows discerning readers to determine whether the inclusion/exclusion criteria have been applied as intended (Schlosser, Wendt, & Sigafoos, 2007), is available upon request.

Data Extraction

Using a specifically developed coding form and manual (Schlosser & Wendt, 2006),[1] the first and second authors independently read and coded each study in terms of the following categories: (a) participant characteristics (number, chronological age, gender, speech before intervention, initial vocal imitation skills), (b) AAC approach used, (c) methods (number of sessions, intervention design, the design used to evaluate the impact on speech production [e.g., single posttreatment probe, multiple pre- and posttreatment probes], type of speech [AAC-mode related, activity-related speech, generalized speech], speech measure [e.g., percentage of correct vocalizations] and technique [e.g., elicitation, observation], interobserver agreement, percentage of interobserver agreement sessions, reliability observer status, treatment integrity, treatment integrity measure, percentage of treatment integrity sessions, and treatment integrity observer status), (d) outcomes (speech outcomes and AAC outcomes), and (e) appraisal level of certainty (i.e., inconclusive, suggestive, preponderant, conclusive). Any disagreements, and there were very few, between the first and second author were resolved through a consensus-building process.

To quantitatively assess the outcomes generated from single-subject experimental designs, the PND was applied (Scruggs, Mastropieri, & Casto, 1987). The PND method requires the calculation of nonoverlap between baseline and successive intervention (or generalization/maintenance) phases by identifying the highest data point in baseline and determining the percentage of data points during intervention (or generalization or maintenance) exceeding this level. PND scores can range from 0% to 100% and can be interpreted using the conventions set by Scruggs, Mastropieri, Cook, and Escobar (1986): A PND greater than 90% is considered highly effective, a PND between 70% and 90% is considered fairly effective, a PND between 50% and 70% is considered of questionable effectiveness, and a PND below 50% reflects unreliable or ineffective treatments. When PND scores were summarized across participants or outcome variables, the recommendation of Scruggs et al. (1986) to use the median was not followed. Because in all cases mean and median scores were either identical or very close, due to the small number of individual PND scores, the mean instead of the median was used.

[1] The coding manual and form are available upon request.

For group designs, two different effect size indices were chosen that are based on the standard mean difference: Cohen's *d* and Hedge's *g*. Cohen's *d* was derived by calculating the difference between the two group means (the control group and the experimental group) divided by the standard deviation for those means (J. Cohen, 1988). When group studies had small sample sizes ($N < 20$), Hedge's *g* was used instead of Cohen's *d* to avoid small sample bias (Lipsey & Wilson, 2001). Hedge's *g* is a simple correction of *d* based on the pooled standard deviation. It was computed by using the square root of the mean square error from the analysis of variance testing for differences between the two groups. For one particular group design that was based on multiple regression procedures (Yoder & Stone, 2006), the difference in variances accounted for by the predictor variables, that is, ΔR^2, served as ES index. Standard mean difference ESs such as *d* and *g* can range from –3.00 to 3.00. J. Cohen (1988) established a widely used convention to interpret the magnitude of ES: an ES below .20 is considered a small effect, an ES of .20 to .50 is a medium effect, an ES of .50 to .80 is important, and anything above .80 is considered a large effect. No specific conventions exist for the interpretation of squared multiple correlation ES. The magnitude of the difference in the variance accounted for by predictor variables gives an impression of how much stronger one predictor is over another. For the *F* test for multiple correlation and regression (MCR) used in Yoder and Layton (1988), the appropriate ES index is f^2 (see J. Cohen, 1977); f^2 for MCR is defined as $f^2 = R^2/(1 - R^2)$, where R^2 is the squared multiple correlation. Cohen's conventions for interpreting f^2 are as follows: ESs of .02, .15, and .35 are considered small, medium, and large, respectively (J. Cohen, 1988).

The appraisal consisted of an evaluation of the design and its implementation, interobserver agreement, and treatment integrity. Depending on how these three issues were addressed in a study, the evidence was appraised as (a) conclusive, (b) preponderant, (c) suggestive, or (d) inconclusive based on a taxonomy developed by Simeonsson and Bailey (1991) and subsequently employed by others (Granlund & Olsson, 1999; Millar et al., 2006; Schlosser & Sigafoos, 2002). A study was rated conclusive if the design provided experimental control, interobserver agreement was reliable, and treatment integrity was solid, allowing the conclusion that the speech outcomes for the participant were undoubtedly the result of the AAC intervention. A study was rated as preponderant if there were minor flaws with respect to the design, interobserver agreement, or treatment integrity, resulting in the conclusion that speech outcomes were more likely than not to have occurred as a result of the AAC intervention, but the evidence was not conclusive. Evidence was rated as suggestive if the study had several minor flaws or failed to provide treatment integrity data, leading to the conclusion that it was plausible, but not certain, that speech outcomes were the result of the AAC intervention. Last, a study was rated as inconclusive if there were significant flaws in the design (regardless of interobserver agreement or treatment integrity) that precluded any conclusions regarding the impact of the AAC intervention on speech production).

Results

Results are reported separately for single-subject experimental designs and group designs. To date, there are no accepted methods available for integrating aggregated data from both types of designs (Kavale, Mathur, Forness, Quinn, & Rutherford, 2000).

Single-Subject Experimental Design Studies

Nine single-subject studies involving a total of 27 participants met the inclusion criteria. Data are summarized in Table 2.

Participant characteristics. The studies included between 1 and 6 participants, with a mean age of 81 months (range = 37–144 months). Of these participants, 85% ($n = 23$) were male, and all but one had no functional speech (i.e., fewer than 10 spoken words). This matches the general trend that more males than females are affected by autism (Fombonne, 2003). The determinations as to whether the apparatus to produce speech was intact were based on anecdotal reports rather than an oromotor assessment. In terms of speech imitation prior to intervention, all but one of the studies did not report any information, and one study did not report on 2 out of 3 of its participants. Those studies that did report on prior speech imitation relied on anecdotal reports rather than on formal testing.

Intervention. Most of the studies evaluated the effectiveness of the Picture Exchange Communication System by itself or in comparison to manual signing. Three studies involved speech-generating devices as part of their intervention; one study examined the effects of enhanced milieu teaching to teach requesting with speech-generating devices, whereas two studies were sorting out the effects of speech output when using speech-generating devices.

Design. Data were extracted from the original studies in terms of the research design used to evaluate (a) AAC acquisition and (b) speech production. For studies evaluating the effectiveness of one intervention, the multiple baseline design or multiple probe design was used most frequently ($n = 5$). For studies comparing two interventions to each other, the adapted alternating treatments design appeared to be used most often ($n = 2$). The multiple pre- and within-treatment probe design was employed most frequently to monitor the effects of AAC intervention on speech production ($n = 4$), followed by the multiple pre-, within-, and post-treatment probe design ($n = 3$). Only one study employed a continuous probe design.

Measures. For speech production, data were extracted for the type of speech measured as well as the dependent variable and technique used. Approximately half of the studies measured only speech that was related to the AAC mode taught in that particular study. For example, if a manual sign was taught for "apple," only speech productions relative to the word *apple* qualified. The other half of the studies counted any speech, regardless of whether it was related to the AAC mode. In one study, both AAC mode-related speech for word approximations and any speech for the number of intelligible words were counted.

In terms of speech measures, there were some consistencies and some differences across studies: Several studies

TABLE 2. Included single-subject experimental studies on the effects of AAC intervention on natural speech production in autism.

	Participant characteristics			Intervention		Design		Measures and outcomes					
Study	Name, CA, gender	Prior speech (technique)	Prior speech imitation	AAC	Conditions (sessions)	Intervention[a]	Speech[b]	Speech type	Speech DV (technique)	Speech Outcomes (PND)	AAC DV	AAC Outcomes (PND)	Appraisal
Charlop-Christy et al. (2002)	Alex, 144; M	NFS (anecdotal)	3-word phrases upon request (anecdotal)	PECS	T1 (PECS Academic, 4) T2 (PECS Play, 4)	MBD across subjects	Multiple PWP probes	AAC mode-related[c]	% word vocalizations (elicitation)	T1 (50)[d] T2 (0)	Eye contact, joint attention, or cooperative play	Training (100), post (100), FU (100)	Suggestive: The speech design rules out threats to internal validity; strong IOA data; TI data are lacking
									% words or approximations (imitation)	T1 (33) T2 (0)			
									MLU (observation)	T1 (0) T2 (50)	Requests or initiations	Training (0), post (100), FU (100)	
	Jake, 44; M	NFS (anecdotal)	Sounds (anecdotal)		T1 (5) T2 (5)	MBD across subjects	Multiple PWP probes	AAC mode-related	% word vocalizations (elicitation)	T1 (60) T2 (40)	Eye contact, joint attention, or cooperative play	Training (100), post (100)	
									% words or approximations (imitation)	T1 (100) T2 (80)			
									MLU (observation)	T1 (60) T2 (40)	Requests or initiations	Training (83), post (80)	
	Kyle, 69; M	NFS (anecdotal)	Sounds (anecdotal)		T1 (9) T2 (9)	MBD across subjects	Multiple PWP probes	AAC mode-related	% word vocalizations (elicitation)	T1 (11) T2 (22)	Eye contact, joint attention, or cooperative play	Training (44), post (100)	
									% words or approximations (imitation)	T1 (11) T2 (44)			
									MLU (observation)	T1 (11) T2 (22)	Requests or initiations	Training (100), post (100)	
Ganz et al. (2007)	Gaspar, 53; M	NFS	Approximations (anecdotal)	PECS	PECS (Phases I–IV)	MBD across subjects	Multiple PW probes	Any speech[e]	# of word vocalizations (imitation)	(8)	Requests (i.e., independent exchanges)	(98)	The speech design is mapped onto the intervention design but fails to establish experimental control; the intervention design establishes experimental control for 2 of the 3 participants; strong IOA data, but TI is not reported.
								AAC mode-related	# word approximations (imitation)	(8)			

(table continues)

TABLE 2 (*continued*).

	Participant characteristics			Intervention		Design		Measures and outcomes					
Study	Name, CA, gender	Prior speech (technique)	Prior speech imitation	AAC	Conditions (sessions)	Intervention[a]	Speech[b]	Speech type	Speech DV (technique)	Speech Outcomes (PND)	AAC DV	AAC Outcomes (PND)	Appraisal
	Leo, 37; M	NFS	NR					Any speech	# of word vocalizations (imitation)	(3)	Requests (i.e., independent exchanges)	(84)	
								AAC mode-related	# word approximations (imitation)	(5)			
	Elise, 61; F	NFS	NR					Any speech	# of word vocalizations (imitation)	(0)	Requests (i.e., independent exchanges)	(90)	
								AAC mode-related	# of word approximations (imitation)	(0)			
Olive et al. (2007)	Mikey, 45; M	No speech	No vocal imitation (anecdotal)	SGD	SGD use via enhanced milieu teaching (14)	MPD across subjects	Multiple PW probes	Any speech	# of words or approximations (child-initiated or elicited)	0	Independent requesting (via SGD)	(88)	Suggestive: TI and IOA strong, minor design flaws because teaching and testing were not separated (however, only unprompted responses were counted).
											Requesting (via SGD or gestures)	(100)	
	Terrence, 66; M	No speech	Imitated "more" (anecdotal)		12	MPD across subjects	Multiple PW probes	Any speech		83	Requesting (via SGD)	(92)	
											Requesting (via SGD or gestures)	(92)	
	Rocky, 48; M	No speech	No imitation (anecdotal)		6	MPD across subjects	Multiple PW probes	Any speech		0	Requesting (via SGD)	(100)	
											Requesting (via SGD or gestures)	(57)	

(table continues)

TABLE 2 (*continued*).

	Participant characteristics			Intervention		Design		Measures and outcomes					
Study	Name, CA, gender	Prior speech (technique)	Prior speech imitation	AAC	Conditions (sessions)	Intervention[a]	Speech[b]	Speech type	Speech DV (technique)	Speech Outcomes (PND)	AAC DV	AAC Outcomes (PND)	Appraisal
Parsons & LaSorte (1993)	# 1, 56; M	NFS (apparatus intact; anecdotal)	NR	SGD (with and without speech output	T1 (Speech, 30) T2 (No speech, 30)	A-B-BC-B-BC/A-BC-B-BC	Continuous probes	Any speech	# words or approximations including echolalia (child-initiated or elicited)	T1 (83) T2 (0)	N/A	N/A	Suggestive: The speech design was mapped on to the intervention design. Order effects were controlled across subjects with an *n* of only 3 (between-group *n* requirements not satisfied); no TI data.
	# 2, 61; M				T1 (30) T2 (30)					T1 (75) T2 (0)	NA	N/A	
	# 3, 68; M				T1 (30) T2 (30)					T1 (100) T2 (8)	N/A	N/A	
	# 4, 74; F				T1 (30) T2 (30)					T1 (100) T2 (0)	N/A	N/A	
	# 5, 79; F				T1 (30) T2 (30)					T1 (50) T2 (0)	N/A	N/A	
	# 6, 72; M				T1 (30) T2 (30)					T1 (92) T2 (0)	N/A	NA	
Schlosser, Sigafoos, et al. (2007)	Avery, 108; F	NFS (apparatus intact; anecdotal)	Imitation index of 0% (formal test)	SGD (with and without speech output)	T1 (Speech, 43) T2 (No speech, 43)	AATD	Multiple PWP probes	AAC mode-related	% vocalizations (elicited)	TI (0) T2 (0)	Requesting	T1 (65) T2 (2)	Preponderant: The speech and intervention design are strong and rule out carryover effects and sequence effects; TI & IOA data are strong; however, instruction did not continue until criterion.
	Greg, 96; M	NFS (apparatus intact; anecdotal)	Imitation index of 0% (formal test)		T1 (37) T2 (37)	AATD	Multiple PWP probes	AAC mode-related	% vocalizations (elicited)	T1 (0) T2 (0)	Requesting	T1 (35) T2 (30)	

(table continues)

TABLE 2 (*continued*).

Study	Participant characteristics: Name, CA, gender	Participant characteristics: Prior speech (technique)	Participant characteristics: Prior speech imitation	Intervention: AAC	Intervention: Conditions (sessions)	Design: Intervention[a]	Design: Speech[b]	Measures and outcomes: Speech type	Measures and outcomes: Speech DV (technique)	Measures and outcomes: Speech Outcomes (PND)	Measures and outcomes: AAC DV	Measures and outcomes: AAC Outcomes (PND)	Appraisal
	Matthew, 120; M	NFS (apparatus intact; anecdotal)	Imitation index of 0% (formal test)		T1 (35) T2 (35)	AATD	Multiple PWP probes	AAC mode-related	% vocalizations (elicited)	T1 (0) T2 (0)	Requesting	T1 (37) T2 (0)	
	Michael, 96; M	Six approximations (apparatus intact)	Imitation index of 60% (formal test)		T1 (42) T2 (42)	AATD	Multiple PWP probes	AAC mode-related	% vocalizations (elicited)	T1 (19) T2 (5)	Requesting	T1 (40) T2 (76)	
	Zachary, 120; M	NFS (apparatus intact; anecdotal)	Imitation index of 0% (formal test)		T1 (46) T2 (46)	AATD	Multiple PWP probes	AAC mode-related	% vocalizations (elicited)	T1 (0) T2 (0)	Requesting	T1 (43) T2 (76)	
Tincani (2004)	Carl, 70; M	NFS (apparatus intact; anecdotal)	Speech (anecdotal)	PECS vs. manual signs	T1 (PECS, 21) T2 (Manual signing, 31)	AATD	Multiple PWP probes	AAC mode-related	#/% of words or approximations (elicited)	T1 (100) T2 (100)	Requesting	T1 (73) T2 (82)	Preponderant: The natural speech design and the intervention design involved noncontinuous data collection. Treatment integrity and reliability were strong.
	Jennifer, 80; F	NFS (apparatus intact; anecdotal)	Words and phrases (anecdotal)		T1 (21) T2 (31)	AATD	Multiple PWP probes	AAC mode-related	#/% of words or approximations (elicited)	T1 (100) T2 (100)	Requesting	T1 (100) T2 (78)	

(table continues)

TABLE 2 (*continued*).

	Participant characteristics			Intervention		Design		Measures and outcomes					
Study	Name, CA, gender	Prior speech (technique)	Prior speech imitation	AAC	Conditions (sessions)	Intervention[a]	Speech[b]	Speech type	Speech DV (technique)	Speech Outcomes (PND)	AAC DV	AAC Outcomes (PND)	Appraisal
Tincani et al. (2006)—Study 1	Damian, 122; M	NFS (apparatus intact; anecdotal)	NR	PECS	36	MBD plus CCD	Multiple PW probes	AAC mode-related	% words; (elicited)	(0)	Requesting	(100)	Suggestive: MBD is across only 2 students, and teaching and testing are confounded. IOA and TI are very good.
									% approximations (elicited)	(11); Phases I-III (0); Phase IV (100)			
	Bob, 141; M	NFS (apparatus intact; anecdotal)	NR		26			AAC mode-related	% words (elicited)	(0)	Requesting	(100)	
									% approximations (elicited)	(0)			
Tincani et al. (2006)—Study 2	Carl, 110; M	NR	NR	PECS (Phase IV)	(11)	ABAB	Multiple PW probes	AAC mode-related	% words (elicited)	(0)	NR	NR	Conclusive: Strong design. IOA and TI are strong as well.
									% approximations (elicited)	(100)	NR	NR	

(table continues)

TABLE 2 (*continued*).

	Participant characteristics			Intervention		Design		Measures and outcomes					
Study	Name, CA, gender	Prior speech (technique)	Prior speech imitation	AAC	Conditions (sessions)	Intervention[a]	Speech[b]	Speech type	Speech DV (technique)	Speech Outcomes (PND)	AAC DV	AAC Outcomes (PND)	Appraisal
Travis (2006)	M. M., 106; M	<10 verbal requests or comments per 10 min sessions	NR	PECS	T1 (Structured, 18)	MBD across behaviors	Multiple PWP probes	Any speech	MLU (observations)	T1 (0)	Requesting	T1 (100), T2 (100)	The MBD design was implemented with only one replication (minor design flaw). In addition, there are no TI data available. IOA data are reported and at adequate levels; however, they are not separated for each of the dependent measures and based on only 10% of sessions. There is also the acknowledged co-intervention that took place in the classroom during nontreatment times and at home, and because it is unclear what exactly was implemented and how well it was carried out, the source of the obtained results is difficult to interpret (fatal design flaw).
					T2 (Unstructured, 18)					T2 (12)	Commenting	T1 (100), T2 (25)	
	N. N., 114; M	Some verbal utterances but often echolalic and repetitive	NR		T1 (18)					T1 (80)	Requesting	T1 (100), T2 (100)	
					T2 (18)					T2 (50)	Commenting	T1 (50), T2 (50)	

Note. AAC = augmentative and alternative communication; CA = chronological age in months; PND = percentage of nonoverlapping data; DV = dependent variable; M = male; NFS = no functional speech; PECS = Picture Exchange Communication System; T1 = Condition 1; T2 = Condition 2; MBD = multiple baseline design; PWP = pre-, within-, and posttreatment speech; FU = follow-up; IOA = interobserver agreement; TI = treatment integrity; MLU = mean length of utterance; NR = not reported; F = female; SGD = speech-generating device; AATD = adapted alternating treatments design; PW = pre- and within-treatment speech design; CCD = changing criterion design.

[a]Intervention design is the design used to evaluate AAC acquisition or feedback. [b]Speech design is the design used to evaluate the effects on speech production. [c]Speech that is related to the AAC mode being taught. [d]Outcomes were attained with novel partners, nontraining settings, and stimuli. [e]Speech that may or may not be related to the AAC mode taught.

monitored complete words as well as word approximations, four studies counted word vocalizations or word approximations as correct, and two studies monitored mean length of utterance. In one study, the definition of "spontaneous utterances" was unclear as to whether approximations or just complete words would count—the phrase "any verbalization" seems to suggest that approximations did qualify.

The speech measurement technique varied across studies. In some studies, speech was elicited through the presentation of an object or graphic symbol in anticipation that the child might say the name of the object. In other studies, speech was modeled for subsequent imitation. Finally, in some studies, the child was observed to monitor child-initiated speech as well as speech that was elicited by the communication partners or the computer.

The most frequently monitored AAC dependent measure related to requesting as a communicative function. One study monitored requests or initiations, making it difficult to discern the effect on either individually. One study measured commenting via the Picture Exchange Communication System. One study also examined eye contact, joint attention, and cooperative play as one dependent measure, creating problems in detecting effects for each of these behaviors individually. In one study, an AAC measure was not employed because the focus was on evaluating the effects of AAC as a feedback mode rather the teaching of the acquisition of an AAC mode (see Table 3).

Outcomes. The speech outcomes ranged from a PND of 0 to a PND of 100. Because the studies varied in terms of the specific treatment evaluated, the speech measure employed, and the technique for "eliciting" speech, it was deemed inappropriate to aggregate these outcomes across studies. Instead, the data from participants within a given study were aggregated across the same dependent measures and techniques (see Table 3). Table 3 indicates that 17 out of 21 mean PND values (or 80.95%) are classifiable in the ineffective range (i.e., below 50) whereas 1 PND mean value was in the fairly effective range (70–90), and only 3 mean PND values are classified in the highly effective range (i.e., above 90). For several dependent measures and studies, there was a considerable amount of variation across participants as indicated by the ranges. Thus, although the mean PND reflects a certain level of effectiveness, one has to keep in mind that still some participants scored in the higher ranges. For other dependent measures and studies, there was very little variation.

Table 4 provides the mean PND values and ranges for the AAC outcomes aggregated across participants within the same studies. These data reveal that the majority of interventions were highly or fairly effective, with only a few interventions that were of questionable effectiveness for some outcomes, or they were ineffective. Those studies with the Picture Exchange Communication System as the independent variable and requesting as the dependent variable were aggregated across studies and yielded a mean PND of 95.2, which is highly effective. None of the other studies could be aggregated across studies. Even though there were three studies using speech-generating devices, their purposes

TABLE 3. Natural speech: PND means and ranges per dependent measure aggregated across participants per study.

Study	*N*	AAC intervention	Dependent measures	PND mean	PND range
Charlop-Christy et al. (2002)	3	PECS	Words elicited—academic	40.3 (I)	11–60
			Words elicited—play	17.0 (I)	0–40
			Words/approximations imitation—academic	48.0 (I)	11–100
			Words/approximations imitation—play	41.3 (I)	0–80
			MLU observation—academic	23.7 (I)	0–60
			MLU observation—play	37.3 (I)	22–50
Ganz et al. (2007)	3	PECS	Words imitation	3.7 (I)	0–8
			Word approximations imitation	4.3 (I)	0–8
Tincani (2004)	2	PECS	Words/approximations elicitation—PECS	100 (H)	100–100
		Manual signs	Words/approximations elicitation—manual sign	100 (H)	100–100
Tincani et al. (2006—Study 1)	2	PECS	Words elicitations	0 (I)	0–0
			Approximations elicitation	5.5 (I)	0–11
Tincani et al. (2006—Study 2)	1	PECS	Words elicitation	0 (I)	0
			Approximations elicitations	100 (H)	100
Travis (2006)	2	PECS	MLU observation—structured	40 (I)	0–80
			MLU observation—unstructured	31 (I)	12–50
Olive et al. (2007)	3	SGD: Enhanced milieu teaching	Words/approximations elicitation or imitation	27.7 (I)	0–83
Parsons & LaSorte (1993)	6	SGD—speech	Words/approximations elicitation or child-initiated—speech	83.3 (F)	50–100
		SGD—no speech	Words/approximations elicitation or child-initiated—no speech	1.3 (I)	0–8
Schlosser, Sigafoos, et al. (2007)	5	SGD—speech	Words/approximations elicitation—speech	3.8 (I)	0–19
		SGD—no speech	Words/approximations elicitation—no speech	1.0 (I)	0–5
Total *N*	27				

Note. I = ineffective (<50); F = fairly effective (70–90); H = highly effective (>90).

TABLE 4. AAC outcomes: PND means and ranges per dependent measure aggregated across participants per study.

Study	*N*	AAC intervention	Dependent measures	PND mean	PND range
Charlop-Christy et al. (2002)	3	PECS	Eye contact, joint attention, or cooperative play—training	81.3 (F)	44–100
			Eye contact, joint attention, or cooperative play—post	100 (H)	100
			Eye contact, joint attention, or cooperative play—FU	100 (H)	100
			Requests or initiations	61 (Q)	0–100
Ganz et al. (2007)	3	PECS	Requests	91.6 (H)	84–98
Tincani (2004)	2	PECS	Requests—PECS	86.5 (F)	73–100
		Manual signs	Requests—manual signs	80 (F)	78–82
Tincani et al. (2006—Study 1)	2	PECS	Requests	100 (H)	100
Tincani et al. (2006—Study 2)	1	PECS	NR	NR	NR
Travis (2006)	2	PECS	Requests—structured	100 (H)	100
			Requests—unstructured	100 (H)	100
			Comments—structured	75 (F)	50–100
			Comments—unstructured	37.5 (I)	25–50
Olive et al. (2007)	3	SGD enhanced milieu teaching	Requests—SGD	93.3 (H)	88–100
			Requests—SGD or gestures	83 (F)	57–100
Parsons & LaSorte (1993)	6	SGD: speech vs. no speech	N/A	N/A	N/A
Schlosser, Sigafoos, et al. (2007)	5	SGD: speech	Requests—speech	44 (I)	35–65
		SGD: no speech	Requests—no speech	36.8 (I)	0–76
Total *N*	27				

Note. Q = questionable (50–70).

varied, from evaluating the effectiveness of enhanced milieu teaching with a device to isolating the effects of speech output. The latter studies had the independent variable but not the dependent variable in common.

Group Design Studies

Two group design studies involving 98 participants met the criteria for inclusion. The extracted data are summarized in Table 5.

Participant characteristics. The participants in the two group studies had a mean age of 60 months and 33 months, respectively. Although the exact levels of speech of included participants were not reported, it can be inferred from the inclusion criteria (fewer than 20 or 25 words) that the participants in the group studies presented with somewhat higher levels of speech than did the participants in the single-subject studies. One study assigned participants a priori into a low and high imitator subgroup based on objective pretesting. The groups of participants in the other study did not differ in terms of 30 plus pretreatment variables, but it is unclear to what degree the participants presented with vocal imitation skills.

Interventions and design. One study compared three different methods for introducing manual signs (sign alone, alternating between sign and speech, and simultaneous communication) with speech alone training. The other study compared the Picture Exchange Communication System with Responsive Education and Prelinguistic Milieu Teaching, which is another prelinguistic approach that does not explicitly introduce AAC modalities. Data were extracted from the original studies in terms of the intervention design and speech design used. In both studies, the randomized control trial was used for both purposes.

Measures. Both studies monitored any speech rather than speech tied to specific AAC modalities introduced. One study recorded any speech with the caveat that it needed to relate to the immediate physical or discourse context. In terms of speech measures, both studies counted only child-initiated spoken words. For example, one study measured the number of different child-initiated spoken words (echolalia or responses were not counted). The other study collected two measures: the number of nonimitative spoken acts and the number of different nonimitative words. These measures accommodated approximations and complete words. As a technique, the investigators in both studies relied on observations rather than elicitation during trials.

Outcomes. Because both studies focused on entirely different AAC interventions, no attempt was made to aggregate the outcomes across studies. However, our own ES statistical analyses were implemented for each study. For Yoder and Layton (1988), no statistical difference between any of the four treatment conditions was found (see Table 5). An f^2 ES index was calculated for the multiple correlation and regression test (see J. Cohen, 1977) to examine whether pretreatment speech imitation predicted speech production outcomes. The f^2 for MCR is defined as $f^2 = R^2/(1 - R^2)$, where R^2 is the squared multiple correlation. With $R^2 = .63$ (from

TABLE 5. Included group designs on the effects of AAC intervention on natural speech production in autism.

	Participant characteristics			Intervention	Design & measures		Speech outcomes		Outcomes statistical analyses (.95 confidence interval)					
Study	*N* (CA)	Speech before intervention	Speech imitation	AAC approach	Design, type, & technique	Dependent measures	Condition A: *M* (*SD*)	Condition B: *M* (*SD*)	Cohen's *d*	*d* lower limit	*d* upper limit	Hedge's *g*	*p*	Appraisal
Yoder & Stone (2006)	38 children (33.6 months)	20 different words or fewer in 3 samples	NR	(1) PECS vs. (2) RPMT	RCT; any speech; observations	# nonimitative spoken acts;	(1T2)[a] 3.6 (4.8)	(2T2) 0.6 (4.8)	.63	−1.53	2.91	.61	.03	Conclusive: Strong design. IOA and TI are strong as well.
							(1T3) 5.5 (3.2)	(2 T3) 5.4 (3.2)	.03	−1.41	1.55	.03	.96	
						# different nonimitative words	(1T2) 2.4 (3.6)	(1T2) 0.6 (3.6)	.50	−1.12	2.21	.49	.04	
							(1T3) 3.1 (2.4)	(1T3) 2.9 (2.4)	.08	−1.00	1.22	.08	.93	
Yoder & Layton (1988)	60 children (60 months)	25 words or fewer based on parent report	A priori: low vs. high imitators (objective testing)	(1) Speech alone, (2) sign alone, (3) alternating, (4) simultaneous	RCT; any speech; observations	# of different child-initiated spoken words (echolalia or responses were not counted)	(1)[b] 1.25 (0.67)	(2) 0.78 (0.62)	.73	−0.04	1.50	.19	.06	Suggestive: strong design; lack of treatment integrity data.
							(1) 1.25 (0.67)	(3) 0.97 (0.64)	.43	−0.33	1.18	.11	.25	
							(1) 1.25 (0.67)	(4) 1.01 (0.78)	.33	−0.42	1.08	.09	.37	
							(2) 0.78 (0.62)	(4) 1.01 (0.78)	−.33	−1.08	0.43	−.09	.38	
							(2) 0.78 (0.62)	(3) 0.97 (0.64)	−.30	−1.05	0.45	−.08	.42	

Note. RPMT = Responsive Education and Prelinguistic Milieu Teaching; RCT = randomized controlled trial.

[a]Refers to the time of testing (T1 = pretest; T2 = intervention test, T3 = second intervention test). [b]Numbers refer to the AAC approaches (e.g., 1 = speech alone, 2 = sign alone).

regression data provided by Yoder & Layton, 1988), an f^2 of 1.70 was obtained, indicating a very strong effect (J. Cohen, 1988). The calculation was repeated and confirmed using the statistical analysis software G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007).

In Yoder and Stone (2006), the Picture Exchange Communication System yielded more nonimitative spoken acts and different words than did Responsive Education and Prelinguistic Milieu Teaching early in the study (first comparison), but these differences disappeared later (second comparison).

Discussion

The purpose of this systematic review was to determine the effects of AAC interventions on speech production in children with autism or PDD-NOS. It is important to realize that none of the included and qualifying studies (i.e., those with greater than zero baselines) reported a decline in speech production as a result of AAC intervention. Thus, the frequently raised concern that the introduction of an AAC mode would hinder speech does not seem supported by these data. At this time, this concern cannot be dismissed in its entirety, however, because it is possible that studies with such negative findings do not get published as often as do studies with positive effects on speech production (Balandin, 2007).

Our outcomes calculations showed that most studies yielded some gains in speech production for most participants. Close inspection of the aggregated data reported in Table 3, however, indicates that most of these gains were of modest magnitude; in fact, their PND values rank in the ineffective range according to the criteria by Scruggs et al. (1986). The aggregated data of three studies, however, indicate gains that are fairly effective (Parsons & LaSorte, 1993; with speech output) or highly effective (Tincani, 2004; Tincani, Crozier, & Alazetta, 2006 [Study 2]). It is interesting to note that the study by Tincani (2004; PND of 100 for both participants) and the first study by Tincani et al. (2006; PND of 0 for both participants) produced conflicting results for the same intervention. In the second study by Tincani et al. (2006), one of the participants of the first study was then submitted to a modified Picture Exchange Communication System intervention whereby vocalizations were reinforced—this is not part of the Frost and Bondy (2002) protocol. Only then was an increase in PND noted for approximations but not for complete words. This raises the possibility of individual difference variables being responsible for these differential outcomes: What is it about the participants in the Tincani et al. (2006) study that required the reinforcement when the participants in Tincani (2004) did not seem to need this to achieve successful speech production? Children with autism and PDD-NOS are known to be a very heterogeneous population. Hence, it is likely that these vast individual differences are also reflected in the pool of included studies and may have contributed to the range of speech outcomes observed across included studies. In order to shed more light on just which individual difference variables may affect speech outcomes, future research needs to do better in assessing and reporting what perceptual, speech, and language skills the children bring to the task. For example, only a few studies reported the participants' vocal imitation skills prior to AAC intervention. With a more thoughtful and systematic assessment and reporting of individual characteristics, more studies can be designed that identify potential predictors of speech production (i.e., Yoder & Stone, 2006).

In two studies, no increases were noted for some of the participants. In Olive et al. (2007), speech production remained the same for 2 of the 3 participants. In Schlosser, Sigafoos, et al. (2007), 5 of the 6 participants remained at zero speech production. It is interesting that these participants also reportedly had 0% vocal imitation skills at the onset of intervention.

Our analyses of the group study by Yoder and Layton (1988) revealed no differences across the four conditions in terms of speech production and confirmed the analysis by J. B. Schwartz and Nye (2006) of the same experiment. This differs from the original study, however, in which the authors concluded that speech alone, simultaneous communication, and alternating between speech and sign were more effective than sign alone. Our finding could be due to low statistical power as there were only 15 participants in each group. With a greater *n,* there would have been greater power and possibly statistical differences across the conditions. Our analysis confirmed the original analyses that pretreatment speech imitation skills are a very strong predictor of subsequent speech production, regardless of the treatment conditions (sign alone, speech alone, simultaneous, or alternating). This highlights the importance of monitoring and reporting individual difference variables.

In the group study by Yoder and Stone (2006), two comparisons were determined as significant, and accordingly, two ESs were meaningful: These are the first comparisons with each speech outcome measure. This revealed medium to important effects in favor of the Picture Exchange Communication System over Responsive Education and Prelinguistic Milieu Teaching in terms of nonimitative spoken communicative acts and the number of different nonimitative words. The magnitude of this effect was yielded within the context of child-initiated (i.e., unprompted) speech rather than speech that was somehow elicited by the experimenters. Being a more stringent measure of speech production, this outcome is indeed noteworthy. Yoder and Stone's (2006) study lends further support for the importance of uncovering child characteristics that might facilitate speech acquisition. Their exploratory analysis showed that the growth rate of the number of different nonimitative words was faster with the Picture Exchange Communication System than with Responsive Education and Prelinguistic Milieu Teaching for children who began treatment with relatively high object exploration. On the other hand, the latter group did better than the Picture Exchange Communication System group in children who began treatment with relatively low object exploration.

Directions for Future Research

In addition to the need to identify predictive child characteristics, several avenues for future research appear fruitful. Because most of the existing research has not been hypothesis driven, it would be advantageous to explore some of the hypotheses proposed in the literature (Blischak et al., 2003; Frost & Bondy, 2002; Skinner, 1957; M. Sundberg et al., 1996). For instance, behaviorists have argued that, on the basis of Skinner's (1957) notion of "automatic reinforcement," the use of manual signing together with spoken words as neutral stimuli when paired with a reinforcer will automatically result in the child producing not only manual signs but also natural speech (M. Sundberg et al., 1996). Frost and Bondy (2002) have claimed this speech-enhancing role of automatic reinforcement for the use of the Picture Exchange Communication System as well. Blischak et al. (2003) developed several other hypotheses to explain why the use of speech-generating devices may facilitate natural speech production. One set of hypotheses centers on the motoric effects of using speech-generating devices. For instance, it is conceivable that once a child has established a degree of automaticity in message selection and production, a reduction in physical demands will be realized, which in turn leads to increases in speech production. A related but different argument places importance on the reduction of pressure to speak through the use of alternative means. This reduction may reduce stress on all systems that must be integrated in order to produce speech and, in turn, result in increased speech production. Another set of hypotheses centers on the acoustic effects associated with the use of speech-generating devices. The immediate output provided by these devices not only increases exposure to speech models (whether or not the partner uses augmented input) to the user but also affords consistent acoustic exposure from one production to the next. This may promote attention to and imitation of the speech signal. The consistent pairing of speech with graphic symbols and their referents has been hypothesized to lead to an increased internal phonologic representation of the spoken word. This, in turn, might facilitate speech production.

Thus, although there is no shortage of hypotheses that individually or collectively explain potential increases in speech production as a result of AAC intervention, these hypotheses have yet to be tested in research. Hypothesis-driven research would have the benefit that the design, the choice of the type of speech measured, and the measurement technique logically ensues from the specific hypothesis. For instance, the "automatic reinforcement hypothesis" would lead to a focus on speech that is related to the AAC mode that is being introduced rather than on generalized speech. Additionally, hypothesis-driven research would reduce some of the variability demonstrated in the methodology across the studies reviewed here and would ultimately facilitate more meaningful aggregation of studies that investigate the same hypothesis.

Future research should also examine whether there is any causal relation between the effectiveness of acquisition of the AAC mode to speech production. A study-by-study visual comparison of Tables 3 and 4 does not seem to reveal a consistent pattern. Some studies with very good acquisition results report very poor speech production results (e.g., Ganz et al., 2007; Olive et al., 2007), and others report excellent speech production results (Tincani, 2004). On the other hand, studies reporting questionable to ineffective AAC mode acquisition for requesting seem to consistently report poor speech production (Charlop-Christy et al., 2002; Schlosser, Sigafoos, et al., 2007). It is clear that there is still much to learn concerning this relationship. An interesting topic to pursue would be to ask what happens to the use of the AAC mode once speech has been acquired for the same referents. Will children continue to use both modes consistent with a multimodal approach to communication; will they, on their own, begin to rely solely on speech and drop the AAC mode; or do they require some sort of instruction in order to do so?

The methodological criteria for inclusion in this review were fairly stringent. Still, this systematic review revealed several areas that future research should more closely observe. One issue relates to the selection and description of participants (Bedrosian, 2003). Here, it would be helpful if the determination of whether a participant presents with "no functional speech" were based on more objective data and operational definitions rather than anecdotal reports. Similarly, in future studies, researchers should objectively assess and report participants' speech imitation skills prior to intervention. This appears to be a promising potential predictor of subsequent speech acquisition as a result of AAC intervention (Yoder & Layton, 1988).

Limitations

An attempt was made to locate unpublished theses and dissertations. This resulted in the inclusion of one thesis. With this yield, a formal analysis of publication bias was not possible. It is unknown whether there are other relevant unpublished papers sitting in researchers' file drawers, because we made no attempt to collect these. Thus, the possibility that the absence of published data of AAC interventions hindering speech is due to publication bias cannot be ruled out with this review effort. Second, although a few non-English articles were considered and ruled out (e.g., Koita & Sonoyama, 2004; Türkbay, Karaman, & Çiyiltepe, 2005), our systematic search strategy was restricted to primarily English databases and journals. Therefore, it cannot be ruled out that our findings are due to a language bias. Finally, although the second author coded all of the studies as well, we do acknowledge that the first author did code one of his own studies.

Conclusions

There is currently no evidence that AAC intervention hinders speech production in children with autism or PDD-NOS. At the same time, this systematic review suggests that the observed gains in speech production may vary across individuals and, if they do occur, are typically small in magnitude. Hence, it is important that families of children with autism and speech-language pathologists are provided with appropriate guidance to develop realistic expectations as far as anticipated gains in speech production. The potential lack of natural speech production gains, as evident in some studies, does not negate the value of AAC interventions. In fact, gains in speech production ought to be viewed as a bonus of AAC interventions rather than as an expectation per se. On the other hand, some children may produce large gains in speech production, but as of yet very little is known about the characteristics of these children.

Whether children with autism are able to imitate speech before starting an AAC intervention is the only known strong predictor of subsequent speech production as a result of manual sign intervention. Future research is needed to identify whether this predictor also applies to other AAC modalities and what other individual difference variables might contribute to speech production. In addition, future research should be hypotheses driven so that the research base can grow in a manner that is conducive to subsequent aggregation.

References

Studies that were included are identified with an asterisk.

American Speech-Language-Hearing Association. (2002). *Augmentative and alternative communication: Knowledge and skills for service delivery* [Knowledge and Skills]. Available from www.asha.org/policy.

Anderson, A. E. (2002). Augmentative communication and autism: A comparison of sign language and the Picture Exchange Communication System (Doctoral dissertation, University of California at Santa Barbara, 2001). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 62,* 4269.

Balandin, S. (2007). Unaided AAC interventions appear to facilitate the development of speech. Rigorous research is required to establish the effects of AAC systems and interventions on speech production across a variety of populations with developmental disabilities [Abstract]. *Evidence-Based Communication Assessment and Intervention, 1,* 63–64. Abstract of Millar, D., Light, J. C., and Schlosser, R. W. (2006). *Journal of Speech, Language, and Hearing Research, 49,* 248–264.

Barlow, D. H., & Hersen, M. (1973). Single case experimental designs: Uses in applied clinical research. *Archives of General Psychiatry, 29,* 319–325.

Bedrosian, J. (2003). On the subject of subject selection in AAC: Implications for planning and interpreting efficacy research. In R. W. Schlosser (Ed.), *The efficacy of augmentative and alternative communication: Towards evidence-based practice* (pp. 57–83). San Diego, CA: Academic Press.

Blischak, D. M., Lombardino, L. J., & Dyson, A. T. (2003). Use of speech-generating devices: In support of natural speech. *Augmentative and Alternative Communication, 19,* 29–35.

Bopp, K. D., Brown, K. E., & Mirenda, P. (2004). Speech-language pathologists' roles in the delivery of positive behavior support for individuals with developmental disabilities. *American Journal of Speech-Language Pathology, 13,* 5–19.

Carr, E. G., Binkoff, J. A., Kologinsky, E., & Eddy, M. (1978). Acquisition of sign language by autistic children. I: Expressive labelling. *Journal of Applied Behavior Analysis, 11,* 489–501.

Carr, D., & Felce, J. (2007a). Brief report: Increase in production of spoken words in some children with autism after PECS teaching to Phase III. *Journal of Autism and Developmental Disorders, 37,* 780–787.

Carr, D., & Felce, J. (2007b). The effects of PECS teaching to Phase III on the communicative interactions between children with autism and their teachers. *Journal of Autism and Developmental Disorders, 37,* 724–737.

Charlop-Christy, M. H., Carpenter, M., Le, L., LeBlanc, L. A., & Kellet, K. (2002). Using the Picture Exchange Communication System (PECS) with children with autism: Assessment of PECS acquisition, speech, social-communicative behavior, and problem behavior. *Journal of Applied Behavior Analysis, 35,* 213–231.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic Press.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, M. (1979, April). *The development of language behavior in an autistic child using a total communication approach*. Paper presented at the Annual International Convention, The Council for Exceptional Children, Dallas, TX.

DiCarlo, C., Stricklin, S., Banajee, M., & Reid, D. (2002). Effects of manual signing on communicative verbalizations by toddlers with and without disabilities in inclusive classrooms. *Journal of the Association for Persons with Severe Handicaps, 26,* 1201–1226.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.

Ferrara, S. M. (2005). Voice output communication aids or picture communication system: The efficient way for individuals with severe disabilities to acquire communication (Master's thesis, California State University, Fullerton). Retrieved February 10, 2007, from ProQuest Digital Dissertations database (Publication No. AAT 1429622).

Fombonne, E. (2003). The prevalence of autism. *Journal of the American Medical Association, 289,* 1–3.

Frost, L., & Bondy, A. (2002). *Picture Exchange Communication System training manual* (2nd ed.). Newark, DE: Pyramid Education Products.

Ganz, J. (2002). Collateral effects of an augmentative communication system on word utterances in children with characteristics of autism (Doctoral dissertation, University of Kansas, 2001). *Dissertation Abstracts International: Section B: The Sciences and Engineering, 63,* 3516.

Ganz, J. B., & Simpson, R. L. (2004). Effects on communicative requesting and speech development of the Picture Exchange Communication System in children with autism. *Journal of Autism and Developmental Disorders, 34,* 395–409.

Ganz, J. B., Simpson, R. L., & Corbin-Newsome, J. (2007). The impact of the Picture Exchange Communication System on requesting and speech development in preschoolers with autism spectrum disorders and similar characteristics [Electronic version]. *Research in Autism Spectrum Disorders, 2,* 157–169.

Goldstein, H. (2002). Communication intervention for children with autism: A review of treatment efficacy. *Journal of Autism and Developmental Disorders, 32,* 373–396.

Granlund, M., & Olsson, C. (1999). Efficacy of communication intervention for presymbolic communicators. *Augmentative and Alternative Communication, 15,* 25–37.

Howlin, P., Gordon, R. K., Pasco, G., Wade, A., & Charman, T. (2007). The effectiveness of picture exchange communication (PECS) training for teachers of children with autism: A pragmatic group randomized controlled trial. *Journal of Child Psychology and Psychiatry, 48,* 473–481.

Kavale, K. A., Mathur, S. R., Forness, S. R., Quinn, M. M., & Rutherford, R. B. (2000). Right reason in the integration of group and single-subject research in behavioral disorders. *Behavioral Disorders, 25,* 142–157.

Koita, H., & Sonoyama, S. (2004). Communication training using the Picture Exchange Communication System (PECS): Case study of a child with autistic disorder. *Japanese Journal of Behavior Analysis, 19,* 161–174.

Koita, H., Sonoyama, S., & Takeuchi, K. (2003). Communication training with the Picture Exchange Communication System (PECS) for children with autistic disorder: The training program and current and future research. *Japanese Journal of Behavior Analysis, 18,* 120–130.

Koul, R., Schlosser, R. W., & Sanscribrian, S. (2001). Effects of symbol, referent, and instructional variables on the acquisition of aided and unaided symbols by individuals with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities, 16,* 162–169.

Lancioni, G. E., O'Reilly, M. F., Cuvo, A. J., Singh, N. N., Sigafoos, J., & Didden, R. (2007). PECS and VOCAs to enable students with developmental disabilities to make requests: An overview of the literature. *Research in Developmental Disabilities, 28,* 468–488.

Layton, T. L. (1988). Language training with autistic children using four different modes of presentation. *Journal of Communication Disorders, 21,* 333–350.

Light, J. C., Beukelman, D. R., & Reichle, J. (2003). *Communicative competence for individuals who use AAC: From research to effective practice*. Baltimore: Brookes.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lloyd, L. L., Fuller, D. R., & Arvidson, H. H. (1997). *Augmentative and alternative communication: A handbook of principles and practices*. Needham Heights, MA: Allyn & Bacon.

Millar, D., Light, J., & Schlosser, R. (2006). The impact of augmentative and alternative communication intervention on speech production of individuals with developmental disabilities: A research review. *Journal of Speech, Language, and Hearing Research, 49,* 248–264.

Mirenda, P. (2001). Autism, augmentative communication, and assistive technology: What do we really know? *Focus on Autism and Other Developmental Disabilities, 16,* 141–151.

Mirenda, P. (2003). Toward functional augmentative and alternative communication for students with autism: Manual signs, graphic symbols, and voice output communication aids. *Language, Speech, and Hearing Service in Schools, 34,* 203–216.

Mirenda, P., & Erickson, K. (2000). Augmentative communication and literacy. In A. Wetherby & B. Prizant (Eds.), *Autism spectrum disorders: A transactional developmental perspective* (pp. 333–367). Baltimore: Brookes.

***Olive, M., de la Cruz, B., Davis, T. N., Chan, J. M., Lang, R. B., O'Reilly, M. F., & Dickson, S. M.** (2007). The effects of enhanced milieu teaching and a voice output communication aid on the requesting of three children with autism. *Journal of Autism and Developmental Disorders, 37,* 1505–1513.

***Parsons, C. L., & LaSorte, D.** (1993). The effect of computers with synthesized speech and no speech on the spontaneous communication of children with autism. *Australian Journal of Human Communication Disorders, 21,* 12–31.

Paul, R. (2003). Promoting social communication in high functioning individuals with autistic spectrum disorders. *Child Adolescent Psychiatry Clinics of North America, 12,* 87–106.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden, MA: Blackwell.

Potter, B., & Brown, D. (1997). Review of the studies examining the nature of selection-based and topography-based verbal behavior. *Analysis of Verbal Behavior, 14,* 85–104.

Schlosser, R. W. (2003). Effects of AAC on natural speech development. In R. W. Schlosser (Ed.), *The efficacy of augmentative and alternative communication: Towards evidence-based practice* (pp. 403–425). New York: Academic Press.

Schlosser, R. W., & Blischak, D. M. (2001). Is there a role for speech output in interventions for persons with autism? A review. *Focus on Autism and Other Developmental Disabilities, 16,* 170–178.

Schlosser, R. W., & Goetze, H. (1992). Effectiveness and treatment validity of interventions addressing self-injurious behavior: From narrative reviews to meta-analysis. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (Vol. 7, pp. 135–175). Greenwich, CT: JAI Press.

Schlosser, R. W., & Sigafoos, J. (2002). Selecting graphic symbols for an initial request lexicon: Integrative review. *Augmentative and Alternative Communication, 18,* 102–123.

***Schlosser, R. W., Sigafoos, J., Luiselli, J., Angermeier, K., Schooley, K., Harasymowyz, U., & Belfiore, J.** (2007). Effects of synthetic speech output on requesting and natural speech production in children with autism. *Research in Autism Spectrum Disorders, 1,* 139–163.

Schlosser, R. W., & Wendt, O. (2006). *The effects of AAC intervention on speech production in autism: A coding manual and form*. Unpublished manuscript, Northeastern University, Boston.

Schlosser, R. W., Wendt, O., Angermeier, K., & Shetty, M. (2005). Searching for and finding evidence in augmentative and alternative communication: Navigating a scattered literature. *Augmentative and Alternative Communication, 21,* 233–255.

Schlosser, R. W., Wendt, O., Bhavnani, S., & Nail-Chiwetalu, B. J. (2006). The use of information seeking strategies in evidence-based practice: The case of pearl growing. *International Journal of Language and Communication Disorders, 41,* 567–582.

Schlosser, R. W., Wendt, O., & Sigafoos, J. (2007). Not all systematic reviews are created equal: Considerations for appraisal. *Evidence-Based Communication Assessment and Intervention, 1,* 138–150.

Schwartz, I. S., Garfinkle, A. N., & Bauer, J. (1998). The Picture Exchange Communication System: Communicative outcomes for young children with disabilities. *Topics in Early Childhood Special Education, 18,* 144–159.

Schwartz, J. B., & Nye, C. (2006). A systematic review, synthesis, and evaluation of the evidence for teaching sign language to children with autism. *EBP Briefs, 1,* 1–17.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research methodology: Methodology and validation. *Remedial and Special Education, 8,* 24–33.

Scruggs, T. E., Mastropieri, M. A., Cook, S. B., & Escobar, C. (1986). Early intervention for children with conduct disorders: A quantitative synthesis of single-subject research. *Behavioral Disorders, 11,* 260–271.

Sigafoos, J., & Drasgow, E. (2001). Conditional use of aided and unaided AAC: A review and clinical case demonstration. *Focus on Autism and Other Developmental Disabilities, 16,* 152–161.

Simeonsson, R., & Bailey, D. (1991). Evaluating programme impact: Levels of certainty. In D. Mitchell & R. Brown (Eds.), *Early intervention studies for young children with special needs* (pp. 280–296). London: Chapman and Hall.

Skinner, B. F. (1957). *Verbal behavior.* Englewood Cliffs, NJ: Prentice Hall.

Son, S.-H. (2006). Comparing two modes of AAC intervention for children with autism (Doctoral dissertation, University of Texas at Austin, 2005). *Dissertation Abstracts International: Section A: Humanities and Social Sciences, 66,* 2544.

Son, S. H., Sigafoos, J., O'Reilly, M., & Lancioni, G. E. (2006). Comparing two types of augmentative and alternative communication for children with autism. *Pediatric Rehabilitation, 9,* 389–395.

Stoner, J. B., Beck, A. R., Bock, J. B., Hickey, K., Kosuwan, K., & Thompson, J. R. (2006). The effectiveness of the Picture Exchange Communication System with nonspeaking adults. *Remedial and Special Education, 27,* 154–165.

Sundberg, C., & Michael, J. (2001). The benefits of Skinner's analysis of verbal behavior for children with autism. *Behavior Modification, 25,* 698–724.

Sundberg, M. (1993). Selecting a response form for nonverbal persons: Facilitated communication, pointing systems or sign language? *Analysis of Verbal Behavior, 11,* 99–116.

Sundberg, M., Michael, J., Partington, J., & Sundberg, C. (1996). The role of automatic reinforcement in early language acquisition. *Analysis of Verbal Behavior, 13,* 21–37.

Tincani, M. (2003). Effects of selection-based versus topography-based communication training on the acquisition of mands by children with autism and multiple disabilities (Doctoral dissertation, Ohio State University). *Dissertation Abstracts International: Section A: Humanities and Social Sciences, 63,* 2506.

***Tincani, M.** (2004). Comparing the Picture Exchange Communication System and sign language training for children with autism. *Focus on Autism and Other Developmental Disabilities, 19,* 152–163.

***Tincani, M., Crozier, S., & Alazetta, L.** (2006). The Picture Exchange Communication System: Effects on manding and speech development for school-aged children with autism. *Education and Training in Developmental Disabilities, 41,* 177–184.

***Travis, J.** (2006). *The effectiveness of the Picture Exchange Communication System (PECS) as an augmentative communication system for children with autism spectrum disorders (ASD): A South African pilot study.* Unpublished master's thesis, University of Cape Town, Cape Town, South Africa.

Türkbay, T., Karaman, D., & Çiyiltepe, M. (2005). Otistik Bozuklukta Resim Degis-Tokusuna Dayali Iletisim Sisteminin Kullanilmasi: Bir Olgu Sunumu [Picture Exchange Communication System use in autistic disorder: Case report]. *Çocuk ve Gençlik Ruh Sagligi Dergisi, 12,* 24–29.

Weitz, C., Dexter, M., & Moore, J. (1997). AAC and children with developmental disabilities. In S. Glennen & D. DeCoste (Eds.), *Handbook of augmentative and alternative communication* (pp. 395–431). San Diego, CA: Singular.

Wendt, O. (2007). The effectiveness of augmentative and alternative communication for individuals with autism spectrum disorders: A systematic review and meta-analysis (Doctoral dissertation, Purdue University, 2006). *Dissertation Abstracts International: Section A: Humanities and Social Sciences, 68.*

White, H. D. (1994). Scientific communication and literature retrieval. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 41–55). New York: Russell Sage Foundation.

***Yoder, P. J., & Layton, T. L.** (1988). Speech following sign language training in autistic children with minimal verbal language. *Journal of Autism and Developmental Disorders, 18,* 217–229.

***Yoder, P. J., & Stone, W. L.** (2006). A randomized comparison of the effect of two prelinguistic communication interventions on the acquisition of spoken communication in preschoolers with ASD. *Journal of Speech, Language, and Hearing Research, 49,* 698–711.

Yokoyama, K., Naoi, N., & Yakamoto, J.-I. (2006). Teaching verbal behavior using the Picture Exchange Communication System (PECS) with children with autism spectrum disorders. *Japanese Journal of Special Education, 43,* 485–503.

Received July 25, 2007
Accepted November 7, 2007
DOI: 10.1044/1058-0360(2008/021)

Contact author: Ralf W. Schlosser, Department of Speech-Language Pathology and Audiology, Northeastern University, 151B Forsyth, Boston, MA 02115. E-mail: r.schlosser@neu.edu.