Purpose: To provide clinicians with evidence-based strategies to facilitate early speech development in young children who are not readily imitating sounds. Relevant populations may include, but are not limited to, children with autism spectrum disorders, childhood apraxia of speech, and late-talking toddlers.

Method: Through multifaceted search procedures, we found experimental support for 6 treatment strategies that have been used to facilitate speech development in young children with developmental disabilities. Each strategy is highlighted within this article through a summary of the underlying rationale(s), empirical support, and specific examples of how it could be applied within intervention.

Conclusions: Given the relatively sparse experimental data focused on facilitating speech in children who do not readily imitate, theoretical support emerges as particularly key and underscores the need for clinicians to consider why they are doing what they are doing. In addition, this review emphasizes the need for the research community to bridge the gap between pressing clinical needs and the limited evidence base that is currently available.

Key Words: speech treatment, toddlers, preschool children, apraxia

A speech-language pathologist sits on the floor across from Micah, a 2-year-old boy with few intelligible words and a limited repertoire of speech sounds. Despite Micah’s speech-language difficulties, there is no direct evidence of any neuromuscular difficulties.

In the clinician’s attempt to engage Micah, she pulls two of his favorite toys out of her bag—little plastic animals. She holds up a dog in one hand and a cat in the other.

“Which one do you want, Micah?” He reaches for the dog.

“You want the dog?” she inquires, stressing the name of the desired object. Micah silently reaches again for the desired toy.

The clinician pulls it just out of his reach and instructs, “Say dog.” Micah is silent.

The clinician prompts him again, this time with a simplified version of the word, “Watch my mouth. Say da.” but before she is able to finish, Micah has wandered off. The clinician is frustrated. Although Micah has a small repertoire of speech sounds, including /d/, which he uses in spontaneous jargon, he will not imitate any of them. How is she supposed to help him talk intelligibly?

Children like Micah, who are not readily imitating sounds despite typical strength and structure of oral neuromusculature, are not uncommon on the caseloads of practicing speech-language pathologists, particularly those in early intervention. Relevant populations may include children with apraxia of speech, autism spectrum disorders (ASD), or those referred to descriptively as “late-talking toddlers.” Although goals for language development are primary in many of these children, speech sound development may also be key. Recent research in the treatment of speech sound disorders has focused primarily on how to select target speech sounds (e.g., Kamhi, 2006) and relied heavily on traditional practices of speech sound elicitation that build on imitation abilities. For example,
Van Riper’s hierarchy of production training begins by attempting to elicit imitation of isolated speech sounds (Van Riper & Erickson, 1996). Similarly, Strand and Skinder (1999) note that integral stimulation, one of the most common approaches to treating childhood speech disorders, requires children to imitate models by the clinician. This void between the literature that emphasizes imitation as a basic strategy for facilitating speech sound production and the skills of young children who do not readily imitate leaves clinicians struggling to determine best practice through a process of trial and error. From this rich experience base, a few clinicians have taken the initiative to share their strategies through books, videos, Web sites, and conference presentations (e.g., Boshart, 2004; Hammer, 2006; Marshalla, 2003). In addition, the marketplace for practicing clinicians has been flooded with materials, such as whistles and chewy tubes, that are advertised as products to facilitate early speech production. Despite the wealth of good ideas embedded in many of these resources, they are rarely grounded within a solid theoretical framework or linked to the small repository of empirical research on treatment efficacy that is available. Consequently, clinicians are left to filter through such products and advice without much guidance in regard to evidence-based practice (EBP).

EBP has been defined as a framework for clinical decision making that integrates scientific evidence with clinical expertise and client values (American Speech-Language-Hearing Association [ASHA], 2004; Dollaghan, 2007; Johnson, 2006). The importance of EBP has recently been emphasized in multiple ASHA publications (e.g., ASHA, 2004, 2005; Mullen, 2007). Although the “gold standard” for evidence of treatment efficacy is empirical data from randomized experimental designs emerging from a variety of independent investigators (ASHA, 2004; Johnson, 2006; Mullen, 2007; Odom et al., 2005), few treatment approaches in speech-language pathology have accumulated this form of support. How then do clinicians answer pressing questions regarding the treatment of children on their caseloads? Fortunately, EBP encompasses varied forms of evidence, including theoretical grounding and clinical expertise, and does not preclude use of experimental interventions. However, EBP dictates that we be aware of what evidence does or does not exist for our practices and that we give thought to why certain strategies and techniques might be successful. Without an understanding of why something works, there is always the danger that it will be misapplied, and we will be left unable to build on the strategy to form new and potentially better ideas.

The purpose of the present article is to highlight strategies for facilitating early speech sound development based on the current evidence base. Although the strategies may apply more broadly, the focus is on young children who are not readily imitating early speech sounds despite typical strength and structure of peripheral oral musculature. The clinical goal is to facilitate speech development through eliciting relevant child vocalizations, whether imitative or spontaneous.

We began the process by generating a list of strategies with an established theoretical framework that were being promoted by prominent clinical authorities in the treatment of speech in children with developmental disabilities (e.g., Hammer, 2006; Manolson, 1992; Marshalla, 2003; Square, 1999; Strand, 1995). We then used multifaceted search procedures to identify relevant intervention studies, including electronic keyword searches of relevant databases (e.g., PsycInfo, ERIC, PubMed) and ancestral searches of references in identified sources. Based on this review, we identified six strategies with Level II empirical support (see ASHA, 2004; Johnson, 2006; Justice & Fey, 2004, for a review of evidence levels). Specifically, we selected strategies that were supported by at least one experimental or quasi-experimental study on a relevant population. Consistent with criteria outlined in Dollaghan (2007) in regard to the evaluation of treatment studies, we considered a study as experimental if it actively manipulated treatment in order to examine its effect on meaningful speech outcomes in a relevant pediatric population using either a comparison group or condition. Due to the relatively sparse literature, we have included semieperimental studies that lack the random group assignment or naive evaluation usually associated with the gold standard in experimental design. To be considered a relevant population, the treatment study had to include children with substantial speech impairment associated with developmental disability. Table 1 lists the six treatment strategies that met our criteria of evidential support. The remaining document serves to summarize the theoretical and empirical support for each of these strategies as well as to underscore important areas for future intervention research.

Before proceeding, we want to emphasize our focus on eliciting speech-like vocalizations in young children who are not readily imitating speech sounds. If and when speech imitation skills emerge, additional strategies for shaping speech productions are likely to apply. In addition, the present strategies do not explicitly address important goals related to vocabulary and pragmatic development that will also be key in many of these children. In sum, this article is not intended as a “cookbook” for how to provide intervention. On the contrary, the focus on theoretical frameworks and the empirical research related to each strategy is intended to encourage clinicians to think independently and critically about how each strategy may or may not apply to an individual case.

Evidence-Based Strategies for Eliciting Speech-Like Vocalizations

1. Provide Access to Augmentative and Alternative Communication (AAC)

Rationale. AAC encompasses the use of manual signs, communication boards/books, high-tech electronic devices, and other forms of unaided and aided communication (Binger & Light, 2006, p. 200). We highlight the use of AAC as the first strategy because of its relatively strong empirical support and its underappreciated role in natural speech development. Although AAC can serve as a compensatory means of facilitating social interactions for children with substantial speech-language difficulties (e.g., Charlop-Christy, Carpenter, Le, LeBlanc, & Kellet, 2002; Cumley & Swanson, 1999), we want to stress its utility in facilitating natural speech development. Although counterintuitive to many, AAC can serve as a critical tool in facilitating speech development.
<table>
<thead>
<tr>
<th>Strategy</th>
<th>Primary rationale</th>
<th>Example</th>
<th>Key empirical evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Provide access to AAC</td>
<td>Successful communication enhances the development of relevant semantic and syntactic networks (i.e., it provides a meaningful framework for production of new words, syllables, and sounds), and in the case of voice output devices, provides a timely auditory model.</td>
<td>Provide a child with a “core vocabulary book”: a small album that includes photos of meaningful objects, people, and places.</td>
<td>Millar et al. (2006)</td>
</tr>
<tr>
<td>2. Minimize pressure to speak</td>
<td>Anxiety and stress can have a negative impact on motor performance.</td>
<td>Utilize puppet play to model target sounds or movements.</td>
<td>Baskett (1996), Kouri (2005)</td>
</tr>
<tr>
<td>3. Imitate the child</td>
<td>Imitation may serve as a model for eliciting imitation itself.</td>
<td>If the child babbles “baba,” repeat this back and use it to begin the song “Baa Baa Black Sheep.”</td>
<td>Field et al. (2001), Snow (1989), Tamis-LeMonda et al. (2001)</td>
</tr>
<tr>
<td>4. Utilize exaggerated intonation and slowed tempo</td>
<td>Neural mechanisms devoted to melodic production can be used to “bootstrap” speech production.</td>
<td>Produce meaningful words and phrases with exaggerated prosody, such as an elongated rising intonation for the word <em>more</em> when used as a request.</td>
<td>Kouri &amp; Winn (2006), Wade (1996)</td>
</tr>
<tr>
<td>5. Augment auditory, visual, tactile, and proprioceptive feedback</td>
<td>Children with speech sound disorders may be less able to capitalize on the sensory feedback typically available.</td>
<td>Enhance visual/tactile feedback by providing a light tap on your own or the child’s lips while modeling production of a /b/ sound.</td>
<td>Bernard-Opitz et al. (1999), Hailem et al. (2008)</td>
</tr>
<tr>
<td>6. Avoid emphasis on nonspeech-like articulator movements: focus on function</td>
<td>Sensory motor control for speech is somewhat distinct from nonspeech oromotor behaviors.</td>
<td>If alveolar speech sounds are being targeted, incorporating tongue clicks into meaningful play would be more appropriate than tongue wagging or protrusion.</td>
<td>Braslin &amp; Casoella (2005)</td>
</tr>
</tbody>
</table>
by enhancing children’s ability to build relevant semantic and syntactic networks, and in the case of voice output devices, by providing a consistent acoustic model of target sounds and words that children can select as needed (see Millar, Light, & Schlosser, 2006).

Evidence. A review of evidence regarding the impact of AAC on speech-language development in individuals with disabilities is provided by Millar et al. (2006). Of the six studies identified by the authors as demonstrating experimental control, 89% of the individual cases demonstrated increases in speech production with the introduction of AAC, and the remaining 11% of the cases showed no change. None of the 27 cases demonstrated a deleterious effect on speech production; however, a study by Yoder and Layton (1988) highlighted the potential importance of providing verbal models in conjunction with sign if the ultimate goal is to facilitate speech production. On a related note, a single-subject design by Iacono and Duncum (1995) found that the use of sign in conjunction with a speech output device was more effective in eliciting word productions from a 2-year-old child with Down syndrome than use of sign alone. In sum, the literature suggests rather convincingly that AAC not only can also serve as a powerful tool in facilitating natural speech development.

Examples. The many issues involved in successfully selecting and implementing AAC are beyond the scope of this article, and readers are referred to the variety of comprehensive resources (e.g., Beukelman & Mirenda, 2005; Light, Beukelman, & Reiche, 2003; Reiche, Beukelman, & Light, 2002). However, we offer a few examples of low- and mid-tech options that have been successfully implemented with young children. For example, meaningful signs can be modeled directly by the examiner or through hand-over-hand facilitation of the child’s hands, if he or she is receptive to this. Blacklin and Crais (1998) stress the importance of beginning with “survival words” that are likely to be motivating and powerful, such as words that facilitate requesting and protesting (see also Bondy & Frost, 1994). Considerations for selecting signs, especially for children with cognitive or motor coordination constraints, include the iconic and motoric complexity of the signs, with symmetrical contact gestures with high iconicity being easiest. Specific examples include modeling the sign for “more” when a child is reaching for another cookie or “open” when a child is struggling to open a container of interest.

Given the possibility that children with substantial articulation difficulties may also experience difficulties coordinating manual communication, AAC devices that utilize picture symbols and voice output provide another potentially helpful modality. Hammer (2006) suggests a “core vocabulary book” that could include photos or symbols of favored objects, people, and activities. A child who is searching for a favorite toy or wondering where Mom is could point to the relevant item or person in his or her book. As another example, options for snacks and mealtime could be pictorially represented on the refrigerator door, so that a child could point to what he or she wanted to eat either as an initiation or at the prompt of a caregiver. Bondy and Frost (1994) provide a concrete hierarchy of steps that can be implemented to help children discover the power of using pictures to communicate.

Voice output devices might also be appropriate for some children. Examples are varied and range from “talking key chains” that offer up to 10 s of voice recording to dynamic displays with word prediction and synthetic speech output. Examples for how to incorporate the simpler devices include encouraging a child’s use of a speech output device to take his or her turn in “reading” reoccurring words or phrases from a repetitive book. A one-button speech output device with a belt clip could be worn throughout the day and used to participate in a common social routine, such as saying hello, or in making a frequent request, such as going to the bathroom. In addition to the texts cited earlier, the following Web sites are useful in understanding the options and issues surrounding AAC: http://aac.unl.edu/, where the University of Nebraska at Lincoln provides a collection of resources and links related to AAC; www.wati.org/, where the Wisconsin Assistive Technology Initiative sponsors a loaning library, training events, and related resources; and www.attainmentcompany.com/xcart/home.php, the Web site for the Attainment Company, which is known for its variety of relatively low-cost AAC systems.

2. Minimize Pressure to Speak

Rationale. In addition to building children’s linguistic networks, AAC may actually facilitate speech in part by reducing pressure to “speak” in the traditional sense. The general strategy of reducing communicative pressure is based on the premise that high-pressure situations can have a negative impact on motor performance in both adults and children across a variety of tasks, including speaking (e.g., Beuter & Duda, 1985; I. M. Blood, Wertz, Blood, Bennett, & Simpson, 1997; Caruso, Chodzko-Zajko, Bidinger, & Sommers, 1994; Hardy, Mullen, & Martin, 2001; Milne, 1970; Tolkkitt & Scherer, 1986). Consider the difference between conversing with a friend versus delivering a public speech, or singing in the shower versus performing a solo. The precise mechanism through which anxiety affects motor performance is complex, although the general response to stress involves activation of the sympathetic nervous system paired with inhibition of the parasympathetic system (G. W. Blood, Blood, Bennett, Simpson, & Susman, 1994; Robertson, 2004; Weber & Smith, 1990). Relevant effects on speech may include disruption of respiratory function, increased muscular tremor, and altered sensitivity of sensory receptors within articulatory regions (Weber & Smith, 1990).

Evidence. Experimental evidence in support of decreasing communicative pressure comes from studies that have compared children’s communicative behavior across conditions in two different populations: children with autism and children late to talk. Specifically, a group design by Baskett (1996) directly compared the amount of vocalizations in 8 children with autism, age 4–6 years, during two conditions: one in which the child and examiner were in the same room together, and one in which the examiner was in a separate room but could be viewed live by the child through a video monitor. Both conditions included a scripted pseudoconversation intended to elicit interaction. The video condition was expected to increase communicative behaviors by decreasing the stress associated with direct interpersonal interaction. In
support of this prediction, the children’s mean vocal behavior coupled with eye gaze toward the examiner was higher in the video condition than in person. The study also provided preliminary support for the idea that using a puppet as a conversational partner might decrease communicative pressure in some children, as evidenced by increased eye gaze and vocal behaviors (see also Bishop, Hartley, & Weir, 1994).

Additional support for the use of lower-pressure techniques comes from a recent study by Kouri (2005), a group design comparing the effect of two treatment approaches, mand-elicited imitation and modeling with auditory bombardment, on word production in 29 late-talking preschoolers. Although the mand-elicitation condition, which utilized mands (e.g., “What do you want?”) and prompted imitations (e.g., “Tell me bubble”), elicited more target words within the treatment setting, the number or percentage of target words generalized to the home setting did not differ between conditions. The study suggests that higher-pressure techniques, such as direct requests for imitation, do not enhance word learning outside the treatment setting to any greater extent than lower-pressure strategies, such as modeling (see also Camarata, Nelson, & Camarata, 1994, for a similar study focused on grammatical targets).

Examples. Attempts to reduce communicative pressure can take a variety of forms, depending in large part on what circumstances an individual child finds stressful. Examples might include avoiding direct requests for imitation, minimizing time pressure, avoiding “test questions,” following the child’s lead, and utilizing familiar interactions and naturalistic settings (e.g., Girolametto, 1988). Additional clinical strategies include simultaneous vocalization (Velleman, 2006) or “vocal contagion” (Marshalla, 2003) in which children vocalize in conjunction with others, thereby creating less focus on their own voice.

We also recommend the use of puppets as interactive partners. Puppet play reduces the potential power differential associated with adult–child interactions and minimizes the pressure of certain social conventions such as eye contact and conversational turn-taking. Children can either interact directly with the puppet or simply observe interactions between the puppet and clinician. Such interactions can be used to model target sounds (e.g., using a bumblebee puppet to model syllable-initial /b/: “Bumblebee, you had better not eat my banana. Beware bumblebee, if you bite my banana I’ll bop you with my bean bag!”), elicit imitations of relevant oral motor movements (e.g., encouraging a reticent puppet to open its mouth and say “ah” while pretending to play doctor), or provide a comfortable context in which to use unfinished carrier phrases (“Let’s tell the bee when to start—ready, set … go!”) to encourage spontaneous production of /go/.

3. Imitate the Child

Rationale. In addition to being a low-pressure strategy, the primary rationale for imitating the child is that it serves as a means to model the skill of imitation itself. In short, being imitated may teach a child to imitate, and imitation is a useful strategy in learning to use spoken language (see Masur & Eichorst, 2002; Schwartz & Leonard, 1985; Snow, 1989). Recent evidence of mirror neurons is also consistent with the importance of developing imitation skills (Nishitani, Avikainen, & Hari, 2004; Rizzolatti & Craighero, 2004).

Mirror neurons fire when an observer sees someone perform a familiar action, and their firing activates relevant motor neurons. A clinician’s imitation of a child’s speech may trigger mirror neuron firing (Oberman & Ramachandran, 2007), thereby serving as a form of involuntary rehearsal on the child’s part. Of interest, neural imaging studies indicate that the mirror neuron system is impaired in children with ASD (Iacoboni & Dapretto, 2006); in fact, Dapretto et al. (2006) found an inverse relationship between symptom severity and mirror neuron activity in their functional magnetic resonance imaging study of 9 boys with ASD.

Whether or not the adult’s example of imitation leads to imitation by the child in the short term, focusing on the child’s spontaneous activity within his or her natural environment is likely to be more engaging for the child (e.g., Lewy & Dawson, 1992), facilitate generalization of any elicited speech sounds (Koegel, Camarata, Koegel, Ben-Tallal, & Smith, 1998), and simplify the speaking task by using a familiar form to elicit the unfamiliar function of imitation (Velleman, 2006).

Evidence. Imitating the child is incorporated as a single thread in several multifaceted therapy programs (e.g., The Hanen Program for Parents or Greenspan’s Floortime approach). The Hanen Program, for instance, stresses the importance of imitation, specifying “one of the best ways to connect with very young children who are just beginning to communicate is to imitate their sounds, actions, facial expressions and words” (Manolson, 1992, p. 18). Despite relatively widespread use of this strategy in conjunction with other techniques, we focus here on evidence to support its use in isolation. For example, Field, Field, Sanders, and Nadel (2001) examined the effect of adult imitation of child behavior in a sample of 20 nonverbal children with autism, age 4–6 years. Specifically the investigators compared children’s behavior across two conditions: one in which the adult imitated a child’s behaviors, including vocalizations, and the other in which the adult was simply asked to play with the child. The children exposed to the imitation condition demonstrated a significant increase in vocalizations across sessions, whereas the children in the play condition did not.

In addition, work by Snow (1989) including 100 children age 14–20 months demonstrated associations between the frequency with which mothers imitated their children and the frequency with which the children used imitation. As maternal imitation increased, spontaneous imitative behavior increased in their children’s speech as well. Although correlation designs do not specify causal influence, Snow used such data to make the point that imitation is a skill that can be learned, rather than a less mutable personality trait. Similarly, a prospective longitudinal study of 40 mother–child dyads by Tamis-LeMonda, Bornstein, and Baumwell (2001) found the extent to which mothers imitated their 13-month-old children predicted the timing of the children’s later child language milestones even after controlling for child language differences at 13 months. For instance, a mother might respond “Ball!” to a child’s vocalization of “ba,” such a response was found to predict the timing of developments including the child’s acquisition of his or her first 50 words.
the use of combinatorial speech, and first references to the past.

Finally, a study by L. B. Leonard et al. (1982) supported the idea of selecting targets from the child’s current productions, which is a key element of imitation (see also Stelo-Gammon, 1998). Leonard et al. found that young children (age 2;8 [years; months] to 3;1) were significantly more likely to produce new words with phonological characteristics they had already acquired than to produce phonologically challenging words. The authors introduced 16 unfamiliar words to 28 children, 14 with normal language development and 14 with specific language impairment. Half of the items were “out-of-phonology” words containing initial consonants or clusters not yet acquired by the child, and half were “in-phonology” words. Both groups showed a clear preference for the in-phonology words. The intervention strategy employed in this study, in which the in-phonology targets presented to the child were determined by the child’s prior productions, could be viewed as a form of delayed imitation of the child, which in turn facilitated the child’s word production: Examples. A clinician’s imitation of the child might include verbal or nonverbal actions. In regard to the latter, imitating a child’s nonverbal actions, especially oral motor movements, could provide a useful springboard for facilitating verbal interaction. For example, if the child yawns, the clinician might pretend to yawn (with an exaggerated vowel-like sound) and wait to see how the child responds. Similarly, if a child were to make a lip-spreading movement (e.g., to smile for a photo or to “show off” something he or she did), the clinician might imitate the action and pair with a relevant speech sound, like a prolonged /i/ for “See!”

Clinicians can also imitate a child’s spontaneous verbalizations; for example, if the child says “baba,” the adult conversational partner might imitate “baba” and then use it as a start for the song “Baa Baa Black Sheep.” As mentioned by Marshalla (2003), this strategy also allows a clinician to assign meaning to vocalizations that are already in a child’s repertoire. For example, the child may regularly vocalize “dih-uh” in his spontaneous jargon. If he happens to say it while pointing to a desired object, the clinician might imitate and assign meaning (which may or may not be there already in the child’s mind) by saying, “Dih-uh. Oh, this one. You want this one,” and then pausing to give the child a chance to respond verbally or otherwise. Once a child is producing recognizable words, incorrect pronunciations need not be imitated. Instead, imitation could serve as an opportunity to model the adult form of the word in conjunction with additional strategies for facilitating language development, such as expansion, which fall outside the scope of this article (see McCauley & Fey, 2006).

4. Utilize Exaggerated Intonation and Slowed Tempo

Rationale. Whether imitating the child or initiating interaction, exaggerated intonation is a commonly employed technique (e.g., J. S. Leonard, 1992; Macaluso-Haynes, 1985). Although the slower tempo that often characterizes phrases with exaggerated intonation may be facilitatory in its own right, we suggest that the primary rationale for this strategy is that neural mechanisms involved in singing can be used to “bootstrap” speech production due to partially distinct but also overlapping neural networks. In a 2003 article, Patel hypothesized that “linguistic and musical syntax share certain syntactic processes (instantiated in overlapping frontal brain areas) that apply over different domain-specific syntactic representations in posterior brain regions” (p. 679). Similarly, research into the neural bases of pitch processing indicates that both right and left cerebral hemispheres, as well as some subcortical structures, are involved in evaluating pitch changes in both speech and music (Baum & Pell, 1999). However, some imaging studies have described distinct patterns of brain activity for speech and singing. Jeffries, Fritz, and Braun (2003) concluded that speech involved relatively greater neural activation in left-hemisphere structures, while singing the same text resulted in a relative increase in right-hemisphere activity. For children in the target populations, a focus on tasks that emphasize right-hemisphere activation could prove to be a valuable compensatory mechanism.

In addition to the idea of neurological bootstrapping, a secondary rationale for modeling exaggerated prosody is that prosodic features may be key in facilitating speech intelligibility. Even children without clear speech sound production may be able to increase their communicative success by utilizing familiar intonational patterns, such as a rising intonation to mark a question versus a falling intonation to mark a comment (see Hargrove, Roetzel, & Hoodin, 1989). In other words, even if exaggerated intonation does not help with the accuracy or fluency of speech production, it may highlight important features of the speech signal for intelligibility.

Evidence. While the neural pathways involved have yet to be labeled conclusively, the use of music in speech therapy has a long history. Clinicians have used singing in therapy for patients with acquired apraxia (Wambaugh & Doyle, 1994), and the fluency-inducing effects of singing are well documented in the stuttering literature (Alm, 2004). In fact, Melodic Intonation Therapy, developed by Albert, Sparks, and Helm (1973) for adults with aphasia, has been frequently applied to young children with speech and language difficulties. The approach focuses on (a) the addition of a melodic line, (b) slowed tempo, and (c) exaggerated rhythm and stress (Wade, 1996).

Although exaggerated prosody has been used in the treatment of adult communication disorders (Hyland & McNeil, 1987), we focus here on its use to facilitate speech production in young children. Wade (1996) conducted a single-subject alternating-treatment design to compare the effects of Melodic Intonation Therapy versus oral motor treatment on the initial consonant production of a 3-year-old girl with apraxia of speech. When the study began, she had been in treatment for a year and evidenced more than a 1-year delay in the areas of expressive language and phonology. A 2-month therapy program targeted phonemes at varying levels of difficulty, classed as “easy,” “intermediate,” or “hard.” When results were compared for Melodic Intonation Therapy and oral motor therapy, the former was found to be significantly more effective for phonemes at all three difficulty levels (see also Helfrich-Miller, 1984). Additional studies with weaker experimental designs have also noted gains in verbal imitation (Krauss & Galloway, 1982) and general intelligibility (Grube, Spiegel, Buchhop, & Lloyd, 1986) as a function of exaggerating intonational patterns.
In the realm of child language, a recent study by Kouri and Winn (2006) compared the use of sung versus spoken stories to facilitate learning of nonsense words by 16 preschool-age children with language difficulties. Although the sung condition did not differentially facilitate word comprehension, it was associated with more unsolicited attempts at imitation of the target forms during the second of the two experimental sessions. In sum, converging evidence supports the use of melody and exaggerated prosody to facilitate children’s attempts and accuracy at speech production.

**Examples.** Exaggerated prosody can be overlaid onto any meaningful word or phrase. For example, the word *hi* could be produced with an elongated falling intonation, or the word *more* could be modeled with a rising intonation to denote a request. Frequent two-syllable phrases could be presented with exaggerated two-tone patterns (e.g., high-low) and later used as an unfinished carrier phrase (e.g., “Alllll … done” or “Light … on”). Phrases might be practiced with or without musical “accompaniment,” such as striking two keys on a keyboard. Similarly, pauses could be inserted into familiar songs to tempt children into completing the anticipated piece (e.g., “…Like a diamond in the …” or “Row row row your …”).

### 5. Augment Auditory, Visual, Tactile, and Proprioceptive Feedback

**Rationale.** Though established speech production is usually highly automatic and accomplished without attention to sensory feedback (Lof, 2007; see also Todorov, 2004), the latter likely plays an important role in establishing new speech behaviors (Clark, Robin, McCullagh, & Schmidt, 2001). Consequently, when imitation of a new speech movement is challenging under natural circumstances, we recommend harnessing enhanced sensory feedback to guide the new movement. The rationale is that children with speech difficulties may be less able to capitalize on sensory feedback due to one or more of the following reasons: (a) the need to focus the bulk of their attention on the movement itself, (b) a lack of sensory receptors or good sensory function in certain feedback domains, such as audition, or (c) the tendency to experience more “sensory” or “motor noise” (Todorov, 2004, pp. 910–911) than typically developing children. In any case, augmented sensory data may then be needed to develop internal models for speech sound production. Internal models refer to the neural mechanisms that associate motoric plans with their sensory consequences (Brass & Heyes, 2005; Iacoboni, 2005; Iacoboni & Wilson, 2006; Kawato, 1999).

**Evidence.** The small literature on the benefits of enhancing sensory input or feedback focuses largely on manipulation of four domains: auditory, visual, tactile, and proprioceptive. In the auditory domain, the Fast ForWord Program (Scientific Learning Corporation, 1997) was developed to provide children with exposure to acoustically modified speech that involves slowed consonant-vowel transitions; we do not review this literature here due to its focus on language outcomes in school-age children who are likely to have already acquired speech imitation skills (Gillam, Loeb, & Friel-Patti, 2001; Gillam et al., 2008; Merzenich et al., 1996; Tallal et al., 1996, 1997). In addition, there is a small literature regarding the use of visual feedback to facilitate speech sound production using real-time spectrographic displays, ultrasound, and electromagnetic articulography in older children and adults (Adler-Bock, Bernhardt, Gick, & Bacsfalvi, 2007; Barry, 1989; Carter & Edwards, 2004; Katz, Bharadwaj, & Carstens, 1999; Katz et al., 2007; Meredith, 2007; Shuster, Ruscello, & Smith, 1992).

We focus here on two treatment studies that have utilized mixed forms of augmented feedback to facilitate speech development in young children with significant developmental speech disabilities. First, a study by Hailpern, Karahalios, Halle, DeThorne, and Coletto (2008) examined the role of computerized contingent auditory and visual feedback on the vocalizations of 2 nonverbal children with autism, age 5 and 7 years, using an adapted within-subject experimental design. Visual feedback included colorful circles, fireworks, or jagged lines that appeared on a computer screen in response to the loudness, pitch, and duration of the child’s voice. Auditory feedback included pitch-shifted replay of the child’s vocalizations, reverberation, or segments of familiar musical tunes. Results found that computerized feedback increased vocalizations significantly in one child and demonstrated a promising trend toward increasing speech-like vocalizations in the other (see also Bernard-Optiz, Sriram, & Sapuan, 1999). Despite its promise, results from this line of research are extremely preliminary. The extent to which results will generalize across children and across contexts is an area for further investigation.

In the tactile and proprioceptive realms, another more established means of providing augmented feedback to facilitate speech is a cuing system referred to as PROMPT, which stands for Prompts for Restructuring Oral Muscular Phonetic Targets (Chumpelik, 1984; Square, 1999). This system uses a different tactile prompt for each English phoneme according to its place, manner, and voicing in conjunction with providing auditory and visual models (Chumpelik, 1984). Most cited research using PROMPT with children has been disseminated outside of peer-reviewed venues and comes from a single set of investigators (www.promptinstitute.com/; see Smit, 2004, as a potential exception). We will draw attention here to a study by Square, Bose, Goshulak, and Hayden (n.d.) that was accessible online. This pilot study utilized a single-subject design to target production accuracy of trained lexical items in 6 boys, all age 4, with substantial language and phonological disabilities. The authors noted gains in target production accuracy as well as in overall communication and social interaction (see also Rogers et al., 2006).

**Examples.** Enhanced feedback has been advocated for many years by prominent clinical authorities (e.g., Boshart, 2004; Marshalla, 2005, 2007), with techniques such as speaking more loudly and facing the child to accentuate the auditory and visual characteristics of target sounds (Ertmer et al., 2002, p. 193). Focusing first on the auditory realm, target words and sounds can be delivered with slight amplification through headphones (e.g., Hodson & Paden, 1991). Low-tech options, such as talking into an echo microphone, a section of PVC pipe, a mailbox, or any chamber that creates an echo, can be used to attract the child’s interest to the targets.

Enhanced visual feedback can be promoted by face-to-face interaction, mirror work, and gestures. Similarly, a puppet with a movable mouth can provide visual input for speech in a way that may be less threatening to children than being asked to attend directly to an adult (see Strategy 2). Currently
available computer programs that provide real-time visual feedback in response to vocalization are limited to specialized instruments, such as the Visi-Pitch IV by KayPentax, which includes voice games designed for children. In the tactile and proprioceptive domains, touch-pressure cues, such as a light tap on the clinician’s or child’s lips while modeling bilabial plosives, have been utilized (Bleile, 2004). Similarly, brushes (e.g., a toothbrush or cotton swab) and snack foods (e.g., frosting or a lollipop on the alveolar ridge) have all been used to draw attention to key articulator locations.

6. Avoid Emphasis on Nonspeech-Like Articulator Movements: Focus on Function

Rationale. One current controversy in the area of speech motor control involves the relative task-specificity of speech production (see Ballard, Granier, & Robin, 2000; G. Weismer, 2006; Ziegler, 2003). One view, referred to as the task-independent view or the hypothesis of “body part coding,” postulates a general sensory motor mechanism that underlies movement regardless of function, thereby suggesting that the neural basis of imitation is body part-specific (e.g., Goldenberg & Karnath, 2006). Lof (2006) reports that some of the most commonly used oral motor activities include blowing, tongue wagging, and smiling, in which a body part such as the lips or tongue is the focus, yet the direct relation to speech sound production is unclear.

Despite the prevalence of such techniques, converging evidence supports a more task-dependent model, in which the sensory motor control of speech is somewhat distinct from nonspeech oral motor behaviors, such as emotional expression and vegetative functions (see G. Weismer, 2006; Ziegler, 2003). Although the same muscles may be employed for speech and nonspeech oral motor tasks, different muscle fibers may be recruited and in markedly different ways depending on the goal of the action (Iacoboni, 2005; Iacoboni & Wilson, 2006; Moore & Ruark, 1996; Ruark & Moore, 1997). In sum, the task-dependent view stresses the importance of helping individuals form an internal model of the target in regard to motor plans and sensory consequences, and emphasizes the goal or purpose of the motor task as critical to its generalization.

Evidence. In regard to experimental evidence, studies related to the effectiveness of nonspeech activities have generally not supported their use for facilitating speech in children without gross neurological deficits (see reviews by Forrest, 2002; Lof, 2006); however, few studies have been published in peer-reviewed journals. As a notable exception, Braislin and Cascella (2005) published a brief research report on the effect of oral motor treatment on the articulation errors of 4 first-grade students. Treatment included gross motor activities, body positioning, and various articulator movements and resistance exercises. The outcome variable, scores on a standardized articulation test, did not improve substantially after 7 weeks of intervention, which included a total of 15 half-hour sessions (see also Gommerman & Hodge, 1995). Although to date most studies of oral motor activities have focused on older children who are already attempting speech sound imitation, we recommend utilizing nonspeech activities judiciously and only when a child is not yet imitating speech sounds. In addition, the nonspeech activity should mimic the position, movement, and function of the target speech sound(s) as closely as possible. One final note: It is important to distinguish between the use of nonspeech oral motor activities, such as blowing bubbles or licking a lollipop, for therapeutic versus motivational purposes. Such activities may serve as motivating materials through which to implement the strategies we discuss, but may not in and of themselves facilitate speech sound production.

Examples. Examples of incorporating relevant and functional nonspeech activities into therapy vary based on the target. For example, if alveolar sounds are being targeted, modeling tongue clicks would be more suitable than tongue protrusion or tongue wagging, particularly if it could be paired with a meaningful activity, such as a sound effect for a clock or knocking on “doors” in a lift-the-flap book. Similarly, if a clinician wanted to target fricatives, attempting to elicit airflow paired with constriction in the vocal tract is likely to be more facilitative than blowing exercises. For example, one might produce bilabial raspberries during vehicle play to simulate a motor noise. Similarly, velars could be targeted through coughing sounds (e.g., make a kangaroo puppet “cough” and then ask it, “Kangaroo, do you have a cold? Let me get you a Kleenex. Here’s a Kleenex, Kangaroo. Can you cough in the Kleenex? Ready, set . . . oh no, Kangaroo is keeping his cough inside. Let’s help him cough”). As a final example, attempts to elicit /m/ may be better facilitated through use of a kazoo than a whistle because use of a kazoo involves bilabial closure and the initiation of voicing (see Marshalla, 2003).

Strategies for Future Research

Although all of the strategies we have reviewed are in need of additional investigation, this section focuses on four promising strategies that have not received direct examination within a treatment study of childhood speech outcomes. We highlight each of these strategies in conjunction with its theoretical rationale in hopes that it will elicit additional experimental investigation. First, a common strategy in the treatment of children with motor speech disorders is utilizing carrier phrases (Macaluso-Haynes, 1985; Strand, 1995; Yoss & Darley, 1974). The rationale is that carrier phrases provide a rich linguistic context for priming relevant semantic (e.g., Ferrand & New, 2004), structural (e.g., Kemp, Lieven, & Tomasello, 2005; Miller & Deevy, 2006; Savage, Lieven, Theakston, & Tomasello, 2003), and prosodic (e.g., Church & Schacter, 1994; Niedenthal, Krauth-Gruber, & Ric, 2004) networks. As an example, the repetitive phrase “Brown bear, brown bear, what do you see?” (Martin & Carle, 1967) offers a syntactic frame that may prime the target word see via its grammatical properties (i.e., the phrase requires a verb). In addition, it offers a context that activates semantic properties (e.g., the concept of a bear elicits properties of animacy), and the prosodic pattern elicits the opportunity to close the intonational phrase or “melody.” Indeed, theories on the role of prosody in language development and disorders propose that a prosodic hierarchy of feet, prosodic words, and phonological phrase determines young children’s abilities to produce strong and weak syllables (Gerken & McGregor, 1998).

Despite widespread use, the effect of using carrier phrases on speech production has not been isolated within a treatment
study, though their use in language treatment has been documented. For example, Bradshaw, Hoffman, and Norris (1998) found that the cloze procedure paired with expansions facilitated children’s language use more than a combined paradigm of direct questions and direct modeling during book reading in two 4-year-old children with language delay (see also Girolametto, Pearce, & Weitzman, 1996, 1997). Future research needs to evaluate the independent effects of carrier phrases on speech production in children from a variety of relevant populations.

A second strategy worthy of study is pairing vocalizations with “analogous” cross-modality movements. For example, Hammer (2006) suggests modeling fricative sounds, such as /s/, by pulling one’s fingers along a string or pairing specific vowel sounds with gross arm movements that mimic lip positioning, such as spreading arms wide for /i/ or putting arms in a circle for /o/. Key to this strategy is utilizing movements that are relatively automatic and analogous to the target vocalization so as not to compete with the attempted speech production for motoric or cognitive resources. We propose that coordinating or synchronizing movements may simplify motor control for the nervous system (seeBeheshti, 1993; Chang & Hammond, 1987; Gettell, 2006; Peters, 1977). To our knowledge, this strategy has not been directly examined in regard to facilitating children’s speech development. However, there is a long tradition in vocal music pedagogy of analogous cross-modality movement as a teaching tool (e.g., Gruhn, 2002; Wis, 1998).

A third strategy for consideration is encouraging vocal play, particularly play that involves articulator movement relevant to speech sound production (see Strategy 6 above), such as turning on and off voicing with a vowel-like production or exploring variations in pitch. The rationale for encouraging vocal play is based on the hypothesized importance of experience and sensory feedback in forming internal models for speech sound production (Max, Guenther, Gracco, Ghosh, & Wallace, 2004; Mussa-Ivaldi, 1999). Once children learn to associate specific motor commands with their sensory consequences, “motor control can be executed in a pure feedforward manner” (Kawato, 1999, p. 718; see also Kent, 1998). Although not evaluated within an experimental treatment study, support for this strategy comes from links between infants’ vocal play and their early speech development (Oller, 1980; Stark, 1980; Stoel-Gammon, 1998). In addition, a number of studies have documented early speech difficulties in children whose vocal exploration was limited by tracheostomy (Hill & Singer, 1990; Kamen & Watson, 1991; Locke & Pearson, 1990; Ross, 1982).

Finally, for children with early metacognitive abilities, incorporating metaphoric devices to solidify emergent speech sound knowledge should receive investigation. By metaphoric devices, we refer to the pairing of specific speech sounds with meaningful environmental sounds, objects, or events (e.g., referring to /s/ as “the snake sound”). As children try to acquire new speech sounds, they are essentially attempting to store, recall, and consistently produce a new form that for the moment carries little meaning. Motor learning theorists speculate that imagery analogies or metaphors might help to relieve the load on abstract working memory during complex motor tasks by allowing individuals to use visual processing skills or episodic memory and limiting reliance on phonological processing (Baddeley & Hitch, 1974; Liao & Masters, 2001; Poolton, Masters, & Maxwell, 2006). This strategy may be particularly powerful for children with limited verbal working memory capacity (S. E. Weismer, Evans, & Hesketh, 1999) or slowed learning ability (Windsor & Hwang, 1999). Additional examples of such “metaphors” for new target sounds include referring to /m/ as the “yummy food” sound, /p/ as the “popcorn” sound, or /u/ as the “monkey” sound (cf. Bleile, 2004; Ertmer et al., 2002; Lindamood & Lindamood, 2003; Mawhinney & McTeague, 2004; Sindrey, 1997).

Concluding Remarks

In sum, many children who do not readily attempt sound imitation pose a significant clinical challenge in regard to targeting speech development. When experimental support is not readily available, clinicians need to give conscious consideration to the rationale for their techniques, and the research community needs to take seriously the challenge to fill current gaps in the clinical knowledge base. Due to the paucity of explicit and comprehensive resources on facilitating speech sound development in young children who do not readily imitate, we considered theoretical frameworks and empirical research findings to emphasize six strategies to guide intervention. With this in mind, the following text “rewrites” the opening vignette using some of the evidence-based strategies outlined in this article.

In the clinician’s attempt to engage Micah, she pulls two of his favorite toys out of her bag—little plastic animals. She holds up a dog in one hand and a cat in the other.

“Which one do you want Micah?” He reaches for the dog.

“You want the dog?” she inquires, stressing the name of the desired object with exaggerated intonation (Strategy 4). Micah silently reaches again for the desired toy.

“You want the dog,” she says as she touches the picture of the dog on Micah’s voice output device (Strategy 1). The clinician imitates Micah’s attempt to engage Micah, she pulls two of his favorite toys out of her bag—little plastic animals. She holds up a dog in one hand and a cat in the other.

When the dog in hand, Micah begins to walk it across the floor while vocalizing “gogogogo.” The clinician imitates Micah (Strategy 3) by saying, “Go go go. Go dog go.” When Micah abandons his dog in search of more toys, the clinician uses exaggerated falling intonation followed by a pause (Strategy 4) to say, “All … done. We’re all … done. Bye-bye dog.”

“Buh-buh,” adds Micah.

Acknowledgments

This project was supported in part by the U.S. Department of Education, Project FOCAL (Grant H325D070061). Thanks to Annie King and Julie Hengst for sharing their expertise in AAC, Ashley Sharer and Lauren Mueller for their frequent trips to the library (online and in person), and Bonnie Johnson, Pamela Hadley, and Pamela Marshall for their feedback on earlier drafts of the manuscript. Finally, thanks to Z for inspiration.
References


DeThorne et al.: Alternatives to Imitation 143


Received December 11, 2007
Revision received May 28, 2008
Accepted September 19, 2008
DOI: 10.1044/1058-0360(2008/07-0090)

Contact author: Laura S. DeThorne, University of Illinois—Speech & Hearing Science, 901 South Sixth Street, Champaign, IL 61820. E-mail: lauras@illinois.edu.