Vocal Training Procedures: A review of behavioral research

Barbara E. Esch, Ph.D.
BCBA-D, CCC-SLP

Speech training
Where to start?

1. Establish preference
   - For people
   - For voices

2. Increase vocalizations
   - Frequency
   - Variability
   - Loudness (dB)

3. Train Verbal Skills
   (discriminated vocalizations)
   - Mand
   - Tact
   - Echoic
   - Intraverbal

Pre-speech learners
- Few vocalizations
- No VB

Early speech learners
- Weak echoics
- Weak vocal mands

Speakers
- Speech under VB control
- But articulation is unclear

Need more vocals to work with
Reinforce current form as mand
Improve form through echoic
Improve form through echoic
Overview

1. Stimulus-stimulus pairing  
2. Vocal variability training  
3. Rapid motor imitation antecedent procedure
Reinforcing value of those speech sounds increases. It sounds “right.” Observes sound. Speech sounds produced.
Stimulus-stimulus pairing (SSP)

**Purpose**
Increase vocalizations
(so these can then come under VB contingencies as mands, tacts, intraverbals, echos, and others)

**Indicators**
Few vocalizations or little variation in vocalizations

---

**Stimulus Stimulus Pairing Procedure**

1. Pair a (neutral) stimulus with an existing reinforcer (either conditioned or unconditioned). As a result, the previous ‘N’ stimulus acquires reinforcing value.
   - Birnbrauer, 1971; Haines, 1977; Steinman, 1968

2. Any response that produces a stimulus that resembles the (previously paired/neutral) stimulus will be automatically reinforced.
   - Skinner, 1957; Vaughan & Michael, 1982

---

**SSP**

**Rationale**

❖ “…a two-stage conditioning history is necessary…” (Sundberg et al., 1996, p. 22)

1. Pair a (neutral) stimulus with an existing reinforcer (either conditioned or unconditioned). As a result, the previous ‘N’ stimulus acquires reinforcing value.
   - Birnbrauer, 1971; Haines, 1977; Steinman, 1968

2. Any response that produces a stimulus that resembles the (previously paired/neutral) stimulus will be automatically reinforced.
   - Skinner, 1957; Vaughan & Michael, 1982
SSP Procedure

- Pair a preferred stimulus with one that is less preferred (or the value is unknown)
- No response required
Vocalization Baseline

Not enough vocalizing overall

Rx training:
- SSP (pairing)
- Vocal variability
- Alternative comm mode

<table>
<thead>
<tr>
<th>30s</th>
<th>1m</th>
<th>30s</th>
<th>2m</th>
<th>30s</th>
<th>3m</th>
<th>16m</th>
<th>30s</th>
<th>17m</th>
<th>30s</th>
<th>18m</th>
<th>30s</th>
<th>19m</th>
<th>30s</th>
<th>20m</th>
<th>30s</th>
<th>21m</th>
<th>30s</th>
<th>22m</th>
<th>30s</th>
<th>23m</th>
<th>30s</th>
<th>24m</th>
<th>30s</th>
<th>25m</th>
<th>30s</th>
<th>26m</th>
<th>30s</th>
<th>27m</th>
<th>30s</th>
<th>28m</th>
<th>30s</th>
<th>29m</th>
<th>30s</th>
<th>30m</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>/a/</td>
<td>/u/</td>
<td>/u/</td>
<td>/a/</td>
<td>/a/</td>
<td>/u/</td>
<td>/u/</td>
<td>/a/</td>
<td>/a/</td>
<td>/u/</td>
<td>/u/</td>
<td>/a/</td>
<td>/a/</td>
<td>/u/</td>
<td>/u/</td>
<td>/a/</td>
<td>/a/</td>
<td>/u/</td>
<td>/u/</td>
<td>/a/</td>
<td>/a/</td>
<td>/u/</td>
<td>/u/</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Vocalizations occurred in 8/60 intervals

Vocalization Baseline = 13%

Vowels Frequency

Consonants Frequency

• Record speech vocalizations that occur during each 30-sec cell
• Mark + for each separate speech vocalization OR transcribe each phonetically

Not enough vocalizing overall

Rx training:
- SSP (pairing)
- Vocal variability
- Alternative communication mode

SSP

Research support

❖ Sundberg et al. (1996): The Role of Automatic Reinforcement in Early Language Acquisition
❖ Premise: Auditory speech stimuli may not function as a reinforcing stimulus for some learners, as evidenced by few, weak, or inconsistent vocal responses that produce these (speech) auditory stimuli
❖ Participants: 1 TD child, 4 preschoolers, severe-to-mod lang delays
❖ Procedure: 15 pairings/min for a few minutes
❖ Results:
  + All children emitted novel vocal responses (so, pairings increased vocalizations of children with strong speech skills and children with weak pre-intervention repertoires)
  + Temporary effects; vocalizations dissipated within minutes
SSP

Research support

- 11 published SSP studies after Sundberg et al. (1996)
  - 8 showed temporary increases in target vocalizations
  - 3 showed no SSP effects on target vocalizations
- Possible variables affecting outcomes
  - Responder variables: Pre-existing speech repertoire (frequency, topographies); difficulty of targets selected
  - Conditioning variables: presentation sequence of the S-S, # pairings overall, # syllables presented per trial, SPA method and items identified as high-priority

<table>
<thead>
<tr>
<th>Age (Number)</th>
<th>Pre-skills</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sundberg et al. (1996)</td>
<td>2-4 yr (3)  Normal (1)  Mod-sev lang delay (4)</td>
<td>Effects with all skill levels</td>
</tr>
<tr>
<td>Yoon &amp; Bennett (2000)</td>
<td>3-4 yr (3)  0-2 vocal play sounds, no VB, severe SSD</td>
<td>Effects with all, but better for those with stranger pre-skills</td>
</tr>
<tr>
<td>Miguel et al. (2002)</td>
<td>3-5 yr (3)  Minimal vocals, no VB, Dx ASD</td>
<td>Effects with 2 of 3, but worse for those with stranger pre-skills</td>
</tr>
<tr>
<td>Esch et al. (2005)</td>
<td>6-8 yr (3)  No or minimal vocals, no VB, Dx ASD</td>
<td>Not able to establish ECH responses because no effects of SSP (Could not replicate Miguel et al. 2002)</td>
</tr>
</tbody>
</table>

SSP: Early studies

SSP: Isolating procedural variables

What procedural variables might produce a more robust effect?
- Interspersed trials of S+ and S−
- Added a “look!” cue to observe/attend
- Added motherese
SSP: Isolating Procedural Variables

Purpose:
Replicate Esch et al. (2009) to demonstrate generality of the enhanced procedure

<table>
<thead>
<tr>
<th>Age (Number)</th>
<th>Pre-skills</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rader et al. (2014)</td>
<td>4-7 yr (3)</td>
<td>Low vocal play No echoics Dx. ASD Chrom disorders (2)</td>
</tr>
</tbody>
</table>

Copyright © Barbara E. Esch and Esch Behavior Consultants, LLC All Rights Reserved
SSP: Isolating procedural variables

Purpose:
Separate responder variables, such as vocal repertoires, from conditioning variables that might influence conditioning voice as a reinforcer

- Frequent preference assessments
- Increase f of pairings
- Eliminate pre-exposure to target auditory stimulus
- Button press response instead of vocal response

| Petursdottir, Carp, Matthies, & Esch (2011) | 3.5-4 yr | (3) | Hi vocal play, strong echoic, tact, echolalia (2) | Lo vocal play, no VR; Dx: ASD |

|  |  |  | Dx: ASD |

This study suggests: When SSP fails to increase vocalizations, it may be a failure to condition auditory stimuli as reinforcing.

But, it remains unclear which variables are responsible for this failure:
- Assess the value of auditory stimuli separate from their effect on vocalizations.
SSP Discussion

- SSP-induced vocalizations may indicate that sound-making is (at least somewhat) automatically reinforced
- AR vocalizations are acquired early (thus, perhaps easier [Rader et al., 2014]); that is, infants emit AR vocalizations before they acquire complex VB
- We may be susceptible to AR provided through parity with vocalizations of our verbal community (Palmer, 1996)

SSP Research support - Summary

- SSP is aimed at increasing the available pool of speech vocalizations.
- SSP is a preliminary procedure. Presumably,
  - First, pairing establishes some sounds as preferred stimuli.
  - Then, when those stimuli are randomly produced, their higher value results in those responses (that produced those preferred stimuli) being selected into the repertoire (AR).
  - When this happens, these vocalizations will have to be brought under COR as functional verbal operants (e.g., mands, tacts, echoics, intraverbals).
- Timing of teaching mands, tacts, etc is critical, because SSP effects are temporary.
- SSP is not yet a reliable procedure, so its clinical value is questionable. It may be that other procedures (e.g., VV, RMIA) would yield faster, more robust clinical results.
Vocal Variability (VV)

Purpose
Increase topographies of vocalizations

Indicators
Little or no variation in vocalizations

Vocalization Baseline

Vocal variability

Rationale
❖ Many children with developmental delays (e.g., ASD) emit infrequent and/or repetitive (i.e., invariant) speech sounds.
❖ Variability is an operant that can, and does, come under COR. (But note upcoming: Peleg, Martin, & Holth, EABA, Sept 2014)
❖ Lag schedules of reinforcement provide COR for variable responding and have been shown to evoke varied verbal responses in children with a diagnosis of autism. (See Lee et al., 2002; Susa & Schlinger, 2012)
Response Class Variability

<table>
<thead>
<tr>
<th>Invariant</th>
<th>Random</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goodbye</td>
<td>See you later</td>
</tr>
<tr>
<td>Goodbye</td>
<td>Bye Bye</td>
</tr>
<tr>
<td>Goodbye</td>
<td>Bye Bye</td>
</tr>
<tr>
<td>See you later</td>
<td>By you later</td>
</tr>
<tr>
<td>Goodbye</td>
<td>Bye Bye</td>
</tr>
<tr>
<td>See you later</td>
<td>By you later</td>
</tr>
<tr>
<td>Bye Bye</td>
<td>Bye Bye</td>
</tr>
<tr>
<td>See you later</td>
<td>By you later</td>
</tr>
<tr>
<td>Bye Bye</td>
<td>Bye Bye</td>
</tr>
<tr>
<td>See you later</td>
<td>By you later</td>
</tr>
</tbody>
</table>

Teaching and Operant Variability

Vocal variability

General procedure

<table>
<thead>
<tr>
<th>Lag 1 schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials</td>
</tr>
<tr>
<td>Base</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>Response</td>
</tr>
<tr>
<td>Initial (Y)</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Diff</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Diff</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Diff</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Diff</td>
</tr>
<tr>
<td>Same</td>
</tr>
<tr>
<td>Diff</td>
</tr>
</tbody>
</table>

Note: Copyright © Barbara E. Esch and Esch Behavior Consultants, LLC All Rights Reserved
Vocal variability
Research support - Non-vocal/verbal responding

Differential reinforcement of
- Novel movements by porpoises (Pryor, Haag, & O'Reilly, 1969)
- Novel block-building forms (Goetz & Baer, 1973)
- Novel button press sequences (Miller & Neuringer, 2000)
- Variable block-building play responses (Napolitano et al., 2010, extending Goetz & Baer, 1973)

Vocal variability
Research support - Lag Schedules

Variability of complex language
- Novel verbal responses to questions (Lee et al., 2002, 2006)
  - What do you like to do? How are you?
- Extension (Susa & Schlinger, 2012)
  - 2 methodological improvements
    - SPA instead of parent report of reinforcers
    - Eliminated Q “what do you like to do” b/c visible stimuli could evoke responses

Vocal variability (in early speech learners with weak vocal skills)
- Novel vocalizations by low-vocal, non-verbal children with autism (Esch et al., 2009)
- Variability defined as different topography or different sequence
Results

- Variable vocalizations increased
- Overall frequency of vocalizations increased
- No increase in novel phonemes

May have inadvertently constrained variability b/c phonemes are too few, even though variability (sequence differences) was reinforced.
“One caveat suggested by this study is the importance of developing a socially significant definition of vocal variability. We defined vocal variability as any vocalization whose phonemes differed in topography (lee, mop) or in sequence (ub, buh) from those uttered in the previous trial. For both children, vocal responses tended to vary within a phonemic class whose response members required little tongue repositioning (e.g., uh, ah, buh, muh). Hence, defining and reinforcing variability solely on the basis of phonemic sequence may have inadvertently constrained other aspects of variability that are needed for further speech learning. That is, although speech variations were strengthened, they were atypical of those required for fluent speech in which rapid tongue, lip, jaw, and laryngeal movements must necessarily occur to produce a variety of different phonemes in coordinated sequences.”

(Esch, Esch, & Love, 2009, p. 77)
Vocal variability
Research support - Lag Schedules

❖ VV extension (Koehler-Platten et al., 2013)
❖ Novel phonemes evoked with 2 of 3 low-vocal children with autism with little to no echoic repertoire
❖ Programmed COR for novel phonemes, not just for varied phonemes
❖ “…increasing variability without expanding the repertoire of phonemes does not prepare the participant for further vocal training.” (p. 81)
Vocal variability

Research support - Summary

❖ Operant variability may be altered
❖ Lag schedules can increase variability and novelty of speech syllables

Rapid motor imitation antecedent (RMIA)

Purpose

Evoke echoic responses

Indicators

Echoics are weak (inconsistent, inaccurate, delayed); i.e., not under strong control of an echoic stimulus.
RMIA
Rationale
❖ Behavioral momentum: Low probability responses (e.g., vocal imitations) can be evoked when preceded by higher probability responses (e.g., non-vocal imitations) (Mace & Belfiore, 1990; Mace et al., 1988; Nevin, 1983)
❖ Generalized imitation is a functional response class, so unreinforced responses in the class can be maintained if some responses in the class are reinforced (Baer et al., 1967; Lovaas et al., 1966)

But...generalized imitation may be confined within topographic subclasses (e.g., gross motor, fine motor, short vocal, long vocal); further, generalized imitation training, without mand contingencies, hasn’t automatically resulted in vocal imitation (Garcia et al., 1971; Poulson et al., 1993; Ross & Green, 2003; Young et al., 1994)
❖ Echoic responses are imitative responses (Skinner, 1957) and, as such, should be susceptible to COR that evoke and maintain other imitative responses

Reinforcement
❖ Infant vocalizations and motor & vocal imitations increased with contingent attention (e.g., tickles, smiles) from parents compared to fixed-time (i.e., NCR) attention [that was provided] during baseline conditions (Poulson & Kyritzis, 1988)
❖ RMIA includes a mand contingency; it’s vocal imitation training where vocal responses are preceded by non-vocal imitation responses, and are followed by the opportunity to emit a vocal mand (programmed for reinforcement)
RMIA Procedure

❖ Prerequisite skills (Taught before starting RMIA procedure; Ross & Greer, 2003)
  ❖ Sit still
  ❖ Make eye contact
  ❖ Follow simple directions
  ❖ Imitate non-vocal motor movements

❖ Present a series of rapid non-vocal imitation models, ending with the target vocal imitation model

Video
“…we think that the procedure acts to join the different afferent and efferent responses of see and do to hear and say and functions as a new higher order operant.”

(Greer, personal communication, July 24, 2014)

Ross & Greer (2003)
- Echoics were evoked, following a series of rapid motor imitations
- The “run-up” imitation models were faded; echoics then occurred under echoic-only control
- Then, echoic-to-mand transfers were taught, resulting in unprompted vocal mands
- Mands were maintained at 3-month follow up probe
### RMIA studies

<table>
<thead>
<tr>
<th>Age (Number)</th>
<th>Pre-skills</th>
<th>Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-7 yr (5)</td>
<td>No spontaneous speech. No echoic VB. No verbal imitation. Dx ASD.</td>
<td>All acquired generalized vocal imitations. 4 of 5 participants maintained mand VB at follow up &amp; at 3-mo probes.</td>
</tr>
<tr>
<td>3-5 yr (2)</td>
<td>No vocal comm skills. No or weak echoic VB. Gen’zd imitation skills (Experiment 1 results).</td>
<td>Mand &amp; tact forms acquired during RMIA and some maintained at 1-mo follow up. Both echoic/indep tacts required fewer trials to mastery than mands (but tact COR not restricted to gen’zd rfcr only).</td>
</tr>
<tr>
<td>4-5 yr (2)</td>
<td>No functional comm. No or weak echoic VB. Gen’zd imitation skills. Dx PDD.</td>
<td>Echoics evoked during RMIA procedure. Contingent social rfct produced greater echoic effects than a FT schedule. Motor imit high in both rfct conditions.</td>
</tr>
</tbody>
</table>

### RMIA

**Research support - Summary**

- Echoics, mands, and tacts have been produced during the RMIA procedure.
- Further research is needed to:
  - Replicate and extend the (few) studies available.
  - Investigate some of the concerns identified in these studies.
  - Influence on echoic acquisition by mand vs tact contingencies during training.
  - Prerequisite skills e.g., how strong must the generalized imitation repertoire be for RMIA benefit to occur.
  - Methodology:
    - Separate various treatment components.
    - # training trials.
    - Would rapid LR vs imitation produce similar results.

**Contact**

besch1@mac.com