Effectiveness of manual gestures in the treatment of /r/ distortions

Jessica Lynch, M.S.
Heather L. Rusiewicz, Ph.D., CCC-SLP
Caterina Staltari M.A., CCC-SLP
Diane L. Williams, Ph.D., CCC-SLP

Department of Speech-Language Pathology
John G. Rangos, Sr. School of Health Sciences
Duquesne University
Pittsburgh, PA

ASHA Convention
Atlanta, GA November 17, 2012
Thank you:

•Department of Speech-Language Pathology
  Duquesne University
•Staff, Participants, Thesis Advisor & Thesis Committee

John G. Rangos, Sr. School of Health Sciences

•Thesis Advisor & Committee

None of the authors have financial interests in this research.
• One of the most common speech sound disorders (SSDs) (Shuster, Ruscello, & Toth, 1995)

• Most common sound children fail to obtain (Ruscello survey, 1995).
  - Treatment not effective for everyone (Shuster & Ruscello, 1995).
  - Variability - 21 different types of /r/ (Ristuccia, C. & McGovern, S., 2009).
  - Lack of visibility

• 4th most frequent consonant (Shriberg & Kwiatkowski, 1983).

• Psychosocial impact (Crowe Hall, 1991; Silverman & Paulus, 1989; Felsenfeld et al., 1994)

• One of the most neglected research areas in speech therapy (Gibbon & Patterson, 2006).
Clinicians wanted “novel, improved therapy approaches” for persistent SSDs (Ruscello, 1995).

Clinicians use hands frequently & successfully during articulation, phonology, and apraxia therapies.

Expert opinion
Case studies (visual biofeedback)
Single subject designs (CAS) using hands: IST, DTTC, IST, ACT, STP, Jordan’s gestures

Psychosocial needs

Underlying Theoretical Bases:
Resource allocation theory (RAT) (McNeil, Odell, & Tsang, 1991)- decrease cognitive load
Dynamic systems theory (DST) (e.g., Thelen & Smith, 1994) – entrainment/coordination across all systems

Triangulation for /r/
Hands: Jordan’s gestures
(Hall & Jordan; Square, 1999)

Helpful for:
– point of constriction/ articulator placement
– movement
– facilitate correct production of /ʧ/, /s/, /ʃ/, /r/
– self-cue

– observed spontaneously gestures during production
– gestures enhance overall sensory awareness
  – using visual, tactile, and kinesthetic (and auditory) input
• Explore use of manual hand movement, mimicking the tongue, to obtain accurate /r/ productions
  – MANUAL MIMICRY (Rusiewicz, 2009).

• Visual-kinesthetic information about tongue:
  – Movement (high/low & front/back)
  – Timing (assist motor plan)
  – Shape/placement (configuration in oral cavity)
  – Tension (tense/lax)
  – Co-articulation

• Visual feedback for client productions
• Limit excessive verbal /auditory feedback
Does manual cueing (i.e., manual mimicry treatment) have a significant effect on the accuracy (auditory perceptual) of vocalic /r/ speech production in a young adult with long standing residual /r/ articulation error as compared to a no treatment condition?
Specific Research Questions

1. Does manual cueing effect \( /r/ \) productions as measured by two expert listeners’ judgments of accuracy and IPA transcriptions?

2. Does manual cueing effect \( /r/ \) production as measured by twenty eight naïve listeners’ judgments of vocalic \( /r/ \) production accuracy?
To assess the influence of manual mimicry cueing on:

(i) listener perceptual accuracy judgments (of correct / incorrect) by:
   – a clinician,
   – 28 naïve listeners, and
   – two expert listeners

(ii) expert clinician IPA transcriptions during vocalic /r/ production probes produced by a college age female with residual /r/ distortions.
Significance

– Fundamental contribution to /r/ articulation treatment at the single subject level
  • Novel treatment approach for /r/
  • Systematic exploration of manual cueing for articulation disorders (specifically /r/)
  • Contribute to
    – EBP literature for practicing clinicians
    – gesture and motor entrainment literature
Research design

- Single subject ABAB treatment withdraw design
- Nine 60 minute treatment sessions over a two month time span
  - (1-2x/wk)
- Each session consisted of two distinct treatment sessions (25 min)
  - Total of eighteen discrete sessions

- IV: manual mimicry treatment
- DV:
  - Listener perception of accuracy (correct/incorrect) by three types of listeners:
    - Clinician
    - Experts
    - Naive Listeners
  - Expert listener broad transcriptions
Methods: Participant

- English speaking Female; 21 y.o.
- Recruited from Duquesne University’s outpatient clinic
- Hx of SSD (/r/ + 3 other speech sounds /s/, /ʧ/, and /ʃ/)
  » Receiving services 2x/wk
  » Receiving learning support for reading comprehension

<table>
<thead>
<tr>
<th>Test Type</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audiometric screening</td>
<td>WFL</td>
</tr>
<tr>
<td>Informal vision screening</td>
<td>WFL</td>
</tr>
<tr>
<td>Informal voice screening</td>
<td>WFL</td>
</tr>
<tr>
<td>Oral mechanism exam</td>
<td>WFL/ subclinical findings</td>
</tr>
<tr>
<td>Receptive Language (PPVT-3)</td>
<td>SS= 95, 37th percentile, age equivalent = 17;2, low average</td>
</tr>
<tr>
<td>Articulation (GFTA-2)</td>
<td>SS= 55, &lt;1 percentile, age equivalent 4;4</td>
</tr>
<tr>
<td>Non-verbal Cognition (TONI-4)</td>
<td>SS= 96, 39th percentile, age equivalent 14;6, average range</td>
</tr>
</tbody>
</table>
Duquesne University Speech Sound Clinic:

- one-way glass
- audio-video Landro recording
- lapel mic 50cm below participant’s lip
- additional wireless microphone headset
  - positioned 3” below lower lip
  - Shure PG30, PG1 wireless transmitter, and PG4 wireless receiver; sampling rate 48kHz

- Audacity recording software wav. files
  - project rate 44,100 Hz
Methods: Stimuli
<table>
<thead>
<tr>
<th>Procedures/Outcome measures *all 596 probes randomized</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A₁BAB:</strong> Sessions 1-3</td>
</tr>
<tr>
<td>25 word level /r/ probes - all positions</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>AB₁AB:</strong> Sessions 4-8</td>
</tr>
<tr>
<td>15 min of manual mimicry treatment</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>ABA₂B:</strong> Sessions 9-12</td>
</tr>
<tr>
<td>50 word level /r/ probes, all positions</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>ABAB₂:</strong> Sessions 13-18</td>
</tr>
<tr>
<td>Identical to B1 treatment phase procedures, 16 word level probes</td>
</tr>
</tbody>
</table>
Clinician perceptual judgment
Entire data set x2

Expert perceptual judgment
Correct/Incorrect 18% data set = 106 randomized data points

IPA transcription
9% data set = 53 randomized data points

Naive listener perceptual judgment
Correct/Incorrect 31% data set = 184 randomized data points

Expert transcription comparison
Three types of listeners

1. Graduate SLP student clinician
2. Experts –
   two faculty members at Duquesne University Speech Sound Production Clinic with 25+ yrs experience combined
3. Naïve listeners- \((n=28\text{, female, } M=21\text{ y.o.})\)
   – Recruited from Duquesne University SLP program
Table 1  *Phase mean, standard deviation, and range for clinician perceptual judgments of percent correct vocalic /r/ productions.*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean (M)</th>
<th>Standard Deviation (SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline 1</td>
<td>36%</td>
<td>0.05</td>
<td>24 to 48</td>
</tr>
<tr>
<td>Treatment 1</td>
<td>63%</td>
<td>0.1</td>
<td>34 to 75</td>
</tr>
<tr>
<td>Baseline 2</td>
<td>48%</td>
<td>0.04</td>
<td>43 to 52</td>
</tr>
<tr>
<td>Treatment 2</td>
<td>75%</td>
<td>0.06</td>
<td>66 to 83</td>
</tr>
</tbody>
</table>

Clinician perceptual judgment mean % correct vocalic /r/ averaged per condition.
Expert listeners’ perceptual vocalic /r/ productions (correct/incorrect) judgment percent correct per session.

- **Baseline 1**: (M=5%, SD=0.06364)
- **Treatment 1**: (M=56%, SD = 0.34)
- **Baseline 2**: (M=55%, SD=0.43)
- **Treatment 2**: M=90%, SD=0.05

**Expert perceptual judgments mean across conditions.**
### Results - Expert Transcription

<table>
<thead>
<tr>
<th>Error type</th>
<th>position</th>
<th>occurrence</th>
<th>% occurrence by either expert</th>
</tr>
</thead>
<tbody>
<tr>
<td>derhotization</td>
<td>Final (VC)</td>
<td>26/53</td>
<td>49%</td>
</tr>
<tr>
<td>Sound substitution w/r</td>
<td>Initial (word)</td>
<td>1/53</td>
<td>2%</td>
</tr>
<tr>
<td>Sound lengthening</td>
<td>All- vowels and /r/</td>
<td>8/53</td>
<td>15%</td>
</tr>
<tr>
<td>Addition of sounds /ə/, /j/, &amp; /h/</td>
<td>all</td>
<td>5/53</td>
<td>9%</td>
</tr>
<tr>
<td>Retraction /ɔ/</td>
<td>initial</td>
<td>1/53</td>
<td>2%</td>
</tr>
<tr>
<td>Wavering/tremor</td>
<td>all</td>
<td>2/53</td>
<td>4%</td>
</tr>
</tbody>
</table>

Broad transcription of fifty-three data points were conducted by two expert listeners (and one graduate student clinician), while listening to blinded randomized DVD recordings of vocalic /r/ probes (no more than three times) in a quiet environment.
Mean of all naïve listeners’ perceptual judgments of randomized vocalic /r/ productions per condition: baseline 1 ($M = 11\%, SD = 13.07$), treatment 1 ($M = 65\%, SD = 13.56$), baseline 2 ($M = 51\%, SD = 19.38$), treatment 2 ($M = 91\%, SD = 6.83$)
Demonstration of causal effect = data across phases show three demonstrations of effect at three separate points in time.

4 steps:
1. demonstrate predictable baseline pattern
2. examine within phase patterns (consistency/predictability)
3. compare data from adjacent phases for evidence of IV manipulation effect (predictable pattern change in DV)
4. integrate information across phases; determine if three demonstrations of effect occurred at three separate points in time to indicate a functional or “causal” relationship

6 variables:
(1) level, (2) trend, and (3) variability within a similar data series and (4) immediacy of effect, (5) degree of overlap, and (6) consistency of data series between conditions.

(Kratochwill et al., 2010).
Level - the mean of the data within a phase

- Clinician
  - Baseline 1
  - Treatment 1
  - Baseline 2
  - Treatment 2

- Experts
  - Baseline 1
  - Treatment 1
  - Baseline 2
  - Treatment 2

- Naïve Listeners
  - Baseline 1
  - Treatment 1
  - Baseline 2
  - Treatment 2

M = 36%  M = 63%  M = 48%  M = 75%
M = 3%  M = 56%  M = 53%  M = 90%
M = 11%  M = 65%  M = 53%  M = 90%
Trend— the slope of the best fitting straight line within a phase

- **Clinician**
- **Experts**
- **Naïve Listeners**
Variability - the range or SD around the best fitting line within phase

- **Clinician**
  - Baseline 1: SD = 0.57
  - Treatment 1: SD = 0.1
  - Baseline 2: SD = 0.04
  - Treatment 2: SD = 0.07

- **Expert**
  - Baseline 1: SD = 0.58
  - Treatment 1: SD = 0.17
  - Baseline 2: SD = 0.08
  - Treatment 2: SD = 0.13

- **Naive**
  - Baseline 1: SD = 0.02
  - Treatment 1: SD = 0.11
  - Baseline 2: SD = 0.09
  - Treatment 2: SD = 0.07
steps:
1. demonstrate predictable baseline pattern
2. examine within phase patterns (consistency/predictability)

Within-phase data examination is conducted to describe observed patterns and predict expected performance, given no change to the independent variable (Kratochwill et al., 2010).
Across phase analyses

steps:
3. compare data from adjacent phases for evidence of IV manipulation effect (predictable pattern change in DV)
4. integrate information across phases; determine if three demonstrations of effect occurred at three separate points in time to indicate a functional or “causal” relationship

Across phase data examination is conducted to document causal relationship inferring that the outcome variable was directly affected by the manipulation of the independent variable (Kratochwill et al., 2010).
Immediacy of effect— the change in level between the last three data points in one condition to the first three data points in the next condition, and rapidity of effect substantiates the inference that a direct effect of IV manipulation on outcome measures.

- Clinician
- Expert
- Naive
Degree of overlap - the portion of data in one condition that overlaps with the previous condition and the smaller the overlap the more indicative of treatment effect.

```
clinician

POD= 22.22%

expert

POD= 44.44%

naive

POD= 22.22%
```
Consistency of data - examining phases of similar conditions with one another (i.e., all baseline phases) for data patterns. The greater the consistency the more likely a causal treatment effect occurred.
Clinician intra-rater perceptual correct/incorrect judgments of vocalic /r/ productions averaged per session by % accuracy.
Inter-judge (expert 1 and expert 2) reliability measure of percent agreement for perceptual judgments (correct/incorrect) of vocalic /r/ productions for each session

LIPP software for inter-rater broad phonetic transcription agreement: 0.90
Mean percent correct vocalic /r/ for the three listener types across condition, average, and standard deviation of all listeners for an identical data set of 106 randomized probes

<table>
<thead>
<tr>
<th></th>
<th>Clinician</th>
<th>Expert listeners</th>
<th>Naïve listeners</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 1</td>
<td>n = 2</td>
<td>n = 28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline 1</td>
<td>0%</td>
<td>5%</td>
<td>14%</td>
<td>6%</td>
<td>7.09</td>
</tr>
<tr>
<td>Treatment 1</td>
<td>56%</td>
<td>56%</td>
<td>66%</td>
<td>59%</td>
<td>5.77</td>
</tr>
<tr>
<td>Baseline 2</td>
<td>29%</td>
<td>54%</td>
<td>54%</td>
<td>45%</td>
<td>14.15</td>
</tr>
<tr>
<td>Treatment 2</td>
<td>86%</td>
<td>90%</td>
<td>87%</td>
<td>88%</td>
<td>2.08</td>
</tr>
</tbody>
</table>

Baseline 1     Treatment 1      Baseline 2       Treatment 2

clinician
experts
naïve listeners
mean
Difficult elements to treating /r/  
- Variability results in spectrum of acceptable approximations  
- Challenge judging accuracy  
- Experience perceiving more subtle variations  
- Ensure accurate and reliable method of measuring outcome success  
  - “undifferentiated lingual gestures” (Gibbon, 1995; 1999)  
  - Tension vs. placement
Discussion

• Word level
• Greater variance learning new motor pattern
• Visual cueing substantiated
• Visual-kinesthetic
  – Speech and manual movement intrinsically motor-based/entrained
    » motoric realm?
    • Finger tapping study (Smith, McFarland, & Weber, 1986)
• Use of naïve listeners
Manual cueing as a visual cue showed promising results for clinical treatment of /r/ with this individual.

Theoretical implications:

- **RAT** – hands to make memories (Cook, Yi Yip, & Goldin-Meadow, 2010)
  - avoid taxing auditory/sensory feedback loops
    - needed for subtle adjustments during internal calibration of correct productions
      » Sensory feedback loop-
        • visual-kinesthetic external feedback
          • (i.e., force/spatiotemporal relationships articulators)

- **DST**- focus on communicative aspects (Iverson & Thelen, 1999)
  - support link between hand and speech systems
    - support across systems
      » multi-sensory inputs capitalizing on synchrony in the motoric realm
        • self-cueing (also PML- internalization)
Limitations

- limited sample size
- potential clinician treatment bias
- limited training at all levels
- greater need for generalizability/maintenance
• Functionally relevant treatment effect on this young adult with an intractable /r/ error, previously treated unsuccessfully
• Promising results for remediating vocalic /r/ with potential for increased efficacy (given future research) for clinicians
• Ramifications for direction of motor entrainment research
QUESTIONS????


Future Research

- Replication studies (word/conversation level; different investigators; much larger sample size with varied age range)
- /r/ production variability in typical populations
- Alternate DV acoustic analysis of formant frequencies F1-F5
- Measure timing and synchronicity of manual movements solely in the motoric realm during accurate vs. Inaccurate productions
- Thorough post-hoc qualitative analysis of this data set
- Descriptive statistical analysis of data set (repeat with more participants)
- Directly measuring tongue tension vs. Placement
- Basic science research into the underlying mechanism of connection between and across these two systems (manual and speech) with consideration to emerging theories that supplement DST
- Examine self-cueing vs. clinician driven cueing for increased generalization and maintenance of treatment effect
Clinician: session 4- tremulous, wavering quality & correct, but low volume; session 7- spontaneous self-correction, share/solicit judgments, self-rehearsal/unison productions “mimicking you makes it easier”; self-report of “ARE” most difficult; speed up productions- vowel distortions; session 8- lengthening and adding schwa; session 9- spontaneous use of hand cue (withdraw), faster rate, increased approximations, generalization finding- friend reports understanding her better “when you do that hand thing”, both correct/incorrect within one production d/t lengthening; session 13- increased self-reliance allowed fading of cues; session 14- increase in intensity (volume) and tongue tension; session 17- more precise repetition of models

Naïve themes: difficulty judging (especially with increased length and/or vowel distortions), techniques used to judge, syllabic vs. word level differences, naturalness (overall sounded unnatural), salient characteristics heard (i.e., elongation, exaggeration, emphasis, volume increase, tremor, vowel distortion, differences among targets [“IRE” most difficult to judge/ “ORE” most difficult for participant/ “EAR” most consistent], louder productions (increased volume and effort) were more accurate.
<p>| /air/ Initial | /air/ Medial | /air/ Final |
| /ar/ Initial  | /ar/ Medial  | /ar/ Final  |
| /or/ Initial  | /or/ Medial  | /or/ Final  |
| /ear/ Initial | /ear/ Medial | /ear/ Final |
| /ire/ Initial | /ire/ Medial | /ire/ Final |
| /er/ Initial  | /er/ Medial  | /er/ Final  |
| <strong>Prevocalic /r/</strong> | /er/ Medial Unstressed | /rl/ Medial &amp; Final |</p>
<table>
<thead>
<tr>
<th>Baseline 1 presentations</th>
<th>Baseline 2 presentations</th>
<th>CVC probes</th>
<th>Generalization probes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet</td>
<td>Wear</td>
<td>Share</td>
<td>Rope</td>
</tr>
<tr>
<td>Organ</td>
<td>Scar</td>
<td>Drink</td>
<td>Racetrack</td>
</tr>
<tr>
<td>Wire</td>
<td>Year</td>
<td>Rugby</td>
<td>Entrance</td>
</tr>
<tr>
<td>Wrap</td>
<td>Berry</td>
<td>Iron</td>
<td>First</td>
</tr>
<tr>
<td>Raw</td>
<td>Artery</td>
<td>Burger</td>
<td>Crust</td>
</tr>
<tr>
<td>Fort</td>
<td></td>
<td>Siren</td>
<td>Front</td>
</tr>
<tr>
<td>Sailor</td>
<td></td>
<td>Cereal</td>
<td>Story</td>
</tr>
<tr>
<td>Prize</td>
<td></td>
<td>Farm</td>
<td>Fries</td>
</tr>
<tr>
<td>Grape</td>
<td></td>
<td>Steer</td>
<td>Pretzel</td>
</tr>
<tr>
<td>Sore</td>
<td></td>
<td>Army</td>
<td>Recline</td>
</tr>
<tr>
<td>Where</td>
<td></td>
<td>Earn</td>
<td></td>
</tr>
<tr>
<td>Finger</td>
<td></td>
<td>Area</td>
<td></td>
</tr>
<tr>
<td>Roast</td>
<td></td>
<td>Refrigerator</td>
<td></td>
</tr>
<tr>
<td>Very</td>
<td></td>
<td>Author</td>
<td></td>
</tr>
<tr>
<td>Bar</td>
<td></td>
<td>Beret</td>
<td></td>
</tr>
<tr>
<td>Tree</td>
<td></td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>Bread</td>
<td></td>
<td>Trap</td>
<td></td>
</tr>
<tr>
<td>Mark</td>
<td></td>
<td>Rye</td>
<td></td>
</tr>
<tr>
<td>Hairy</td>
<td></td>
<td>Hungry</td>
<td></td>
</tr>
<tr>
<td>Perch</td>
<td></td>
<td>Swear</td>
<td></td>
</tr>
</tbody>
</table>