Sound Production Treatment for Acquired Apraxia of Speech

Julie Wambaugh

VA Salt Lake City Health Care System and the University of Utah
Salt Lake City, UT

Abstract

Acquired apraxia of speech (AOS) has been demonstrated to be responsive to various types of treatments, even when AOS is chronic (Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006a). Of the different types of AOS treatments, the Academy of Neurologic Communication Disorders and Sciences (ANCDS) AOS Treatment Guidelines report identified articulatory-kinematic treatments as having the strongest evidence supporting their use (Wambaugh et al., 2006a & b). Articulatory-kinematic treatments are those that employ techniques to improve speech production through the facilitation of movement, positioning, timing, and coordination of the articulators. Sound Production Treatment (SPT; Wambaugh, Kalinyak-Fliszar, West, & Doyle, 1998) is an articulatory-kinematic treatment that has received relatively more systematic study than other AOS treatments to date. This report describes SPT and the research supporting its application with speakers with AOS.

In 2006, the Academy of Neurologic Communication Disorders and Sciences (ANCDS) Apraxia of Speech (AOS) Treatment Guidelines Committee published a systematic review of the AOS treatment literature along with guidelines for selecting treatments (Wambaugh, Duffy, McNeil, Robin, & Rogers, 2006a, b). The ANCDS review indicated that various AOS treatments can be expected to result in improved speech production in persons with AOS, even when AOS is chronic. Of the 59 reports that comprised the evidence base for the guidelines report, the majority described treatments that were considered to be articulatory-kinematic in nature. That is, the preponderance of the available published evidence concerned treatments designed to improve speech production by improving the movements of the articulators (e.g., timing, spatial targeting, and coordination). The current report describes an articulatory-kinematic treatment for AOS, termed Sound Production Treatment (SPT; Wambaugh, Kalinyak-Fliszar, West, & Doyle, 1998) that has received relatively more systematic study than other AOS treatments to date.

Description of SPT

SPT was developed in the mid-1990s to promote improved articulation of specific sounds targeted for treatment. Sounds selected for treatment are those that are produced erroneously on a consistent basis during pretreatment testing. The level of production at which treatment is directed is individually determined and has included monosyllabic words, multisyllabic words, phrases, and sentences. Thus, SPT may be applied with AOS speakers of various severities. Treatment is conducted using real words whenever possible. SPT combines modeling, repetition, minimal pair contrast, integral stimulation, articulatory placement cueing, and verbal feedback in a response-contingent hierarchy.

The original SPT hierarchy (Wambaugh et al., 1998) has undergone modification in order to capitalize on SPT research findings, accommodate application at different levels of production, and incorporate principles of motor learning. The most recent version of the SPT hierarchy is presented in the Appendix.

A set of 8-10 treatment items is selected based upon the individual AOS speaker's sound production errors. These items are exemplars of the target sound and reflect a variety of phonetic contexts; specifically, a variety of vowels and other non-target consonants are selected. For example, if the problematic sound is word-initial /s/ and production is disrupted at a monosyllabic level, the target items may be sun, seat, sad, song, soup, sick, same, sell. The hierarchy is applied to one treatment item at a time with subsequent steps of the hierarchy being used only upon production of an error. The order of items is randomized, and as many trials as possible of the entire set of items is completed within a treatment session. Typically, four to eight trials of the entire set are possible in a 50- to 60-minute session.

The set of treatment items is used in every treatment session until mastery is reached as evidenced by performance on probes. Treatment is continued until a predetermined criterion of performance is reached (e.g., 80-90% accuracy) over a pre-established number of probe sessions (e.g., three of four consecutive probe sessions). Performance in therapy is not used to determine termination of treatment. Wambaugh, Doyle, Kalinyak, and West (1996) found that treatment performance may not accurately reflect performance when therapy is not being administered. Performance during probes (i.e., the therapist elicits production of the target items without providing therapy or feedback) is considered a more stringent measure of the effects of treatment than measures obtained during treatment. Probe performance with SPT always has been measured at least 1 day following the previous treatment session, and such measurement is recommended for clinical application.

Rationale for Development and Use of SPT

The steps included in the SPT hierarchy were derived from the early AOS treatment research by Rosenbek, Lemme, Ahern, Harris, and Wertz (1973) and reflected those investigators' suggestions that "therapy should concentrate on the disordered articulation...(and) emphasize the regaining of adequate points of articulation and the sequencing of articulatory gestures" (p. 463).

Although the various techniques employed with SPT have been in existence for many years, they are consistent with current models of speech production. For example, in Guenther and colleagues' neural model of speech production, "directions into velocities of articulators" (DIVA; Guenther, 2006), feed forward commands for speech sounds become tuned by the feedback subsystem in repeated production attempts. Auditory models are considered important, as is somatosensory feedback, for refining sound target regions in the DIVA model.

SPT includes repeated practice and opportunities for contrastive practice by the speaker, which may assist the speaker in developing and refining his/her sound target regions. In addition, the steps of integral stimulation (i.e., "watch me, listen to me, say it with me") and articulatory placement instruction allow the therapist to supplement the speaker's own auditory and somatosensory feedback, which also should facilitate the instantiation of feedforward commands. The aspects of modeling/repetition and integral stimulation are also consistent with the recently posited roles of mirror neurons in sound production (Liuzzi et al., 2008; Skipper, van Wassenhove, Nusbaum, & Small, 2007).

Evidence From SPT Investigations

The initial research with SPT (Wambaugh, Kalinyak-Fliszar, West, & Doyle, 1998) was designed to examine treatment effects in terms of acquisition of trained sounds and response

generalization to untrained sounds. Specifically, the research was intended to evaluate whether trained items (i.e., sounds in words) were produced accurately during probes when no treatment was being administered. Probes in this investigation, as in all of the SPT research, consisted of eliciting production of the experimental items via repetition. Items were always administered in random order and no feedback concerning accuracy or production was provided. Of additional interest was whether improvements in production of one sound would generalize to (a) untrained exemplars of that sound, and (b) untrained, different sounds. It was hypothesized that treatment effects might be of a general nature (i.e., not sound specific), and overall speech motor programming skills might improve.

Wambaugh and colleagues (1998) used a multiple baseline design across participants and behaviors with 3 participants with chronic, moderate, or severe AOS. The participants were similar with respect to aphasia type and severity. For each participant, three sounds were trained sequentially so that treatment effects could be determined to be sound-specific or generalized. Results revealed moderate to large effect sizes for all trained sounds in trained and untrained words as measured in probes conducted during the treatment phase (note: probes were conducted at least 1 day following treatment and reflected short term maintenance). For the majority of trained sounds (seven of nine total), longer-term maintenance effects of treatment were strong at 6 weeks follow-up. However, for 2 of the participants, gains were not maintained for one of three trained sounds. Generalization to untrained exemplars of trained sounds was reported, with increases being similar to trained exemplars. Treatment effects did not extend to untrained sounds for 2 of the speakers. However, limited generalization to untrained sounds was noted for 1 speaker, with affricate training being associated with slight improvement in untrained fricatives.

Response generalization to untrained sounds was the focus of a subsequent investigation by Wambaugh and Cort (1998). As in the previous investigation, the participant presented with moderate-severe AOS and had significant Broca's aphasia. The target sounds in this investigation were closely related, as was the nature of the erroneous productions. That is, the target sounds were all voiced stop and affricate consonants. The participant demonstrated difficulty in voicing control; voiced stop and affricate consonants were produced with longer than normal voice onset times (VOTs) and were perceived as voiceless consonants. Two sounds, /b/ and /dʒ/, were trained sequentially, and then /d/ and /g/ were trained simultaneously in the context of a multiple baseline design across behaviors. Treatment resulted in improved production of trained sounds in both trained and untrained words. In addition, treatment resulted in improved production of untrained stops. More specifically, during /b/ training, /d/ and /g/ improved, with effect sizes of 1.6 and 3.6, respectively. During treatment of /dʒ/, /d/ and /g/ evidenced additional improvement but did not reach criteria established for termination of treatment. Consequently, treatment was applied simultaneously to /b/ and /d/, and high, stable levels of responding were reached. The findings from this investigations suggested that SPT may result in generalized improvement of untrained sounds, if (a) error sounds are related in terms of production or (b) if the nature of the sound errors are similar. An interesting secondary finding of this investigation concerned the voiceless cognates of trained sounds. Prior to treatment, voiceless stop and affricate consonants were produced with 99-100% accuracy. Following application of treatment to the voiced stops/affricate, perceptual accuracy of the voiceless cognates fell to 75-93% accuracy, with VOTs being substantially reduced. This finding indicated that unwanted generalization effects also may result from treatment.

Substantial overgeneralization was reported by Wambaugh, Martinez, McNeil, and Rogers (1999). A participant with moderate-severe AOS and Broca's aphasia received treatment applied sequentially to three sounds. Following application of treatment to the first sound, high levels of accuracy were achieved. However, following application of treatment to the second sound, accuracy of production decreased dramatically for the first sound. This pattern of

positive acquisition, followed by disrupted maintenance, also was seen for the second trained sound when a third sound was trained. Post-hoc analysis of the errors revealed that overgeneralization of the sounds under training accounted for the disrupted maintenance. Fortunately, "booster treatment" involving retraining of the previously trained sound(s) and the most recently trained sound resulted in return of accuracy of performance that was then maintained at follow-up. The results of this investigation implied that training of multiple sounds may promote stimulus discrimination and may be important for thwarting unwanted overgeneralization.

Stimulus generalization (i.e., the production of trained behaviors in untrained contexts) is an important treatment outcome that has received extremely limited study in the AOS treatment literature (Wambaugh et al., 2006b). A few SPT investigations have explored stimulus generalization. Wambaugh et al. (1998) measured production of target items in the untrained context of phrase repetition. Results were mixed across participants, with 1 participant showing generalized production for all trained sounds, 1 participant showing improved production in phrases for one trained sound, and 1 participant showing no generalization to phrase production. Positive stimulus generalization findings appeared to be associated with stronger acquisition results.

Wambaugh (2004) measured the stimulus generalization effects of SPT across several untrained contexts with 2 speakers with AOS and aphasia. Both speakers demonstrated strong acquisition and response generalization effects (untrained exemplars of trained sounds). Stimulus generalization results varied within and across participants and appeared to be related to the nature of the stimuli. Treatment was sequentially modified and extended to additional treatment contexts, which resulted in further gains.

Wambaugh and Nessler (2004) also extended treatment to an additional stimulus context with a speaker with moderate-severe AOS and reported positive results. Groups of sounds were trained simultaneously rather than individually, in an effort to avoid overgeneralization and to promote efficiency of treatment. Although positive acquisition effects were noted for all of the trained sounds, generalization to a story completion context was positive for only one sound. Generalization was associated with strong acquisition effects and a long treatment period of high levels of accuracy of production. When treatment was extended to the story completion context, correct productions increased for all sounds in probes of that context. These findings, along with those of Wambaugh (2004), indicate that generalization may occur to additional contexts for some speakers and for some trained sounds without specific instruction in those contexts. When generalization is absent, extension of treatment to different contexts is likely to promote improved productions in those contexts.

Wambaugh and Nessler (2004) also found that training of multiple sounds simultaneously did not appear to be detrimental to the effects of SPT. This finding was in keeping with an earlier investigation in which sounds groups, rather than individual sounds, were trained with a participant with moderate AOS (Wambaugh, West, & Doyle, 1998). Wambaugh et al. (1998) applied a modification of SPT to the treatment of stops, fricatives, and glides/liquids in the context of sentences containing multiple exemplars of those sounds. Results were positive for all trained sentences, as well as for untrained sentences containing the targeted sounds.

The most recent investigation of SPT (Wambaugh & Mauszycki, in press) extended findings to an individual with severe AOS, significant nonfluent aphasia, and verbal perseverations. Treatment was applied sequentially in the context of a multiple baseline design to two sets of items (monosyllabic words), with three consonants targeted in each set. A third phase of treatment entailed training of all target sounds. Findings were consistent with previous SPT investigations in that positive acquisition and response generalization (to untrained exemplars of trained sounds) were observed with no generalization to different,

untrained sounds. Follow-up probing was conducted at 10 and 15 weeks post-treatment. Surprisingly, maintenance effects were strong at 10 weeks, but diminished considerably at 15 weeks, suggesting that longer term maintenance evaluation may be warranted in future investigations.

Summary

In summary, SPT has consistently resulted in improved accuracy of articulation of trained sounds in trained and untrained words. Eight to ten exemplars of the target sound appear to be sufficient to promote response generalization to untrained exemplars of trained sounds. Generalization to untrained sounds that are unrelated (i.e., are not cognates or closely related in manner of production) is likely to be minimal with SPT. Partial generalization has been observed for sounds that demonstrate similar manner of production with similar types of errors. Stimulus generalization has not been evaluated extensively, but may occur to longer and more complex production conditions for some speakers (Wambaugh, 2004; Wambaugh et al., 1998; Wambaugh & Nessler, 2004).

Maintenance effects generally have been strong at 1-2 months post-treatment for sounds that were observed to reach high levels of accuracy of production during treatment. It appears that relatively high levels of accuracy (e.g., 80% correct or greater) should be achieved in treatment in order for gains to be maintained. This suggests that termination of treatment ideally should be determined by performance criteria rather than number of sessions. The findings from the most recent SPT investigation indicate that longer-term periods of post-treatment maintenance measurement may be necessary for both clinical and research purposes.

SPT has been successfully modified for application to multiple sound targets (Wambaugh & Nessler, 2004), sentences containing numerous targets (Wambaugh, West, and Doyle, 1998), and words elicited through sentence completion (Wambaugh & Nessler, 2004). Although there are many aspects of SPT that require additional research, it can be expected that SPT will result in improved production of targeted sounds for speakers AOS.

Julie Wambaugh is a research career scientist at the VA Salt Lake City Health Care System and an associate professor in the Department of Communication Sciences and Disorders at the University of Utah. She teaches graduate courses and conducts research in the area of adult, acquired neurologic communication disorders. She is currently the Coordinator for ASHA's Special Interest Division 2.

References

Guenther, F.H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders, 39,* 350-365.

Liuzzi, G., Ellger, T., Flöel, A., Breitenstein, C., Jansen, A., & Knecht, S. (2008). Walking the talk—Speech activates the leg motor cortex. *Neuropsychologia, 46*(11), 2824-2830.

Rosenbek, J. C., Lemme, M. L., Ahern, M. B., Harris, E. H., & Wertz, R. T. (1973). A treatment for apraxia of speech in adults. *Journal of Speech and Hearing Disorders, 38*, 462-472.

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex, 17*(10), 2387-2399.

Wambaugh, J. L. (2004). Stimulus generalization effects of Sound Production Treatment for apraxia of speech. *Journal of Medical Speech Language Pathology, 12*(2), 77-97.

Wambaugh, J. L., & Cort, R. C. (1998, January). *Treatment for AOS: Perceptual and VOT changes in sound production*. Poster session presented at the biennial Motor Speech Conference, Tucson, AZ.

Wambaugh, J. L., Doyle, P. J., Kalinyak, M. M., & West, J. E. (1996). A minimal contrast treatment for apraxia of speech. *Clinical Aphasiology,* 97-108.

Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. (2006a). Treatment guidelines for acquired apraxia of speech: A synthesis and evaluation of the evidence. *Journal of Medical Speech Language Pathology, 14*(2), xv-xxxiii.

Wambaugh, J. L., Duffy, J. R., McNeil, M. R., Robin, D. A., & Rogers, M. (2006b). Treatment guidelines for acquired apraxia of speech: Treatment descriptions and recommendations. *Journal of Medical Speech Language Pathology, 14*(2), xxxv-ixvii.

Wambaugh, J. L., Kalinyak-Fliszar, M. M., West, J. E., & Doyle, P. J. (1998). Effects of treatment for sound errors in apraxia of speech and aphasia, *Journal of Speech, Language, Hearing Research, 41*, 725-743.

Wambaugh, J. L., Martinez, A. L., McNeil, M. R., & Rogers, M. A. (1999). Sound Production Treatment for apraxia of speech: Overgeneralization and maintenance effects. *Aphasiology, 13*(9-11), 821-837.

Wambaugh, J. L., & Nessler, C. (2004). Modification of sound production treatment for apraxia of speech: Acquisition and generalization effects. *Aphasiology, 18,* 407-427.

Wambaugh, J. L., & Mauszycki, S. (in press). Sound Production Treatment: Application with severe aphasia. *Aphasiology.*

Wambaugh, J. L., West, J. E., & Doyle, P. J. (1998). Treatment for apraxia of speech: Effects of targeting sound groups. *Aphasiology, 12*(7/8), 731-743.

Appendix

SPT Treatment Hierarchy

1) The therapist says the **target item** and requests a repetition (e.g., "say ***sun***").
   a) If correct, the therapist requests additional repetitions (5 times) and then goes to Step 5.
   b) If incorrect, the therapist gives feedback and says: "Now, let's try a different word" and **presents minimal pair item****. (e.g., "that's not quite right, let's try a different word…say ***ton***".
      - If correct, the therapist gives feedback and says: "Now, let's go back to the other word" and goes to Step 2 with the target word.
      - If incorrect, the therapist gives feedback, attempts production with integral stimulation up to 3 times and goes to Step 2 with the target word.
2) The therapist shows the **printed letter** representing the target sound, says the target word, and requests a repetition (e.g., "let's focus on this sound on the card…say ***sun***").
   a) If correct, the therapist requests additional repetitions (5 times*) and goes to the next item.
   b) If incorrect, the therapist goes to Step 3.
3) The therapist uses **integral stimulation** to elicit the target word – "watch me, listen to me, and say it with me" up to 3 times.
   a) If correct, the therapist requests additional repetitions (5 times*) and go to next item.
   b) If incorrect, the therapist goes to Step 4.
4) The therapist gives **articulatory placement cues**, and requests production of the target word again after cueing using integral stimulation (note: cues are dependent upon the errors produced by the client).
   a) If correct, the therapist requests additional repetitions (5 times*).
   b) If incorrect, the therapist goes to Step 5.
5) Go to the **next item.**

***** The hierarchy is response-contingent (steps are only used as needed) and does not reverse directions.

**Will not be used with sentence level stimuli.