PSHA JOURNAL

ANNUAL PUBLICATION
OF THE
PENNSYLVANIA SPEECH-LANGUAGE-HEARING ASSOCIATION

DECEMBER 2013
EDITION

JOURNAL

IS AN ANNUAL PUBLICATION OF THE PENNSYLVANIA SPEECH-LANGUAGE-HEARING ASSOCIATION

700 McKnight Park Drive, Suite 708
Pittsburgh, PA 15237
412.366.9858 (Phone)
412.366.8804 (Fax)
PSHA@psha.org

THE PENNSYLVANIA SPEECH-LANGUAGE-HEARING ASSOCIATION (PSHA), FOUNDED IN 1960, IS A PROFESSIONAL ORGANIZATION OF SPEECH-LANGUAGE PATHOLOGISTS, AUDIOLOGISTS, AND TEACHERS OF THE HEARING-IMPAIRED. AS SUCH, PSHA:

Serves as a liaison between the American Speech-Language-Hearing Association and its Pennsylvania members.
Represents its' members interests in legislative affairs.
Recommends standards for training and practices.
Provides information about effective services and programs and services in communication disorders and other related fields.
Works to inform the public about the professions, careers, programs, and services in the field of communication disorders.
Encourages basic scientific study of the process of individual human communication.
Monitors state Licensure Board activities.

JOURNAL

EDITOR

REVIEW CONSULTANTS

JOURNAL
2013
CONTENTS

EXAMINING THE CLINICAL APPLICATION OF INTRA-ORAL TACTILE BIOFEEDBACK IN SHORT-DURATION THERAPY TARGETING MISARTICULATED /s/

Gordy Rogers
Articulate Technologies, Inc.
Megan S. Chesin
Columbia University

ABSTRACT

This randomized, controlled, single-blind study examined the clinical utility of a tactile biofeedback device to teach correct tongue placement for the /s/ phoneme. 15 school-aged children who misarticulated /s/ were randomly assigned to an experimental group or a control group and treated with eight individual therapy sessions. The experimental group, which used tactile biofeedback via the test article, recorded a statistically significant remediation response ($p < .05$), whereas the control group, which used only traditional phonetic-based treatment, did not show a statistically significant treatment benefit. While these results suggest the test article would be a valuable clinical tool, further research is required to establish efficacy of this approach and whether these results may be replicated in larger-scale studies and non-neurotypical subject populations.

KEY WORDS

Articulation, Speech Sound Disorders,
Tactile Biofeedback, Treatment Study

INTRODUCTION

Articulation and speech sound disorders affect as many as 7.5% of the school-age population (Shriberg & Kwiatkowski, 1994) and can negatively impact teacher perceptions of students with reduced speech intelligibility (Overby, Carrell & Bernthal, 2007) as well as inter-peer relationships among school-age children (Crowe Hall, 1991). Apart from this documented personal cost, articulation and speech sound disorders contribute to an estimated annual cost to society of between $30 billion and $154 billion in lost productivity, special education services, and medical care (Rubens, 2000). In addition, Jacoby, Lee, Kummer, Levin and Creaghead (2002) found that the various treatment methodologies in use today resulted in no measurable progress in remediating speech sound disorders for approximately 28% of the 234 pre-school and school-age children they analyzed. Given the size of this treatment-resistant cohort and the documented social and societal costs of speech sound disorders, it is incumbent upon researchers and clinical practitioners in the field to not only continue to develop improved treatment methods, but also to test these methods through rigorous treatment studies.

The literature in articulation disorders has a relatively extensive record supporting the use of traditional methods of articulation therapy (see Klein, 1996). Phonological and linguistic approaches to treatment have also shown promise in numerous treatment studies (e.g. Broen & Westman, 1990; Gierut, 1998; Major & Bernhardt, 1998; Almost & Rosenbaum, 2000; Rvachew & Nowak, 2001; Tyler & Lewis, 2005). However, therapy that implements one or more of these evidence-based treatment approaches should not only the overall comparative efficacy of a given approach but also the overall cost-effectiveness of a given approach (Gibbard, Coghlan & MacDonald, 2004). Reducing the amount of direct clinician-to-client required to remediate speech sound disorders, as is the aim of a short duration treatment regimen, would in turn reduce the overall cost of therapy (Gibbard, 1994). The field of speech-language pathology has relatively few studies (e.g. Dickson, Marhsall, Boyle, O'Hare, McCarthey & Forbes, 2009; Eiserman, McCoun & Escobar, 1990) that have focused on reducing treatment duration. The current study aims to add to the research corpus focused on short-duration treatment studies.

Jacoby et al. (2002) found that the mean number of 15-minute treatment units required to achieve one level of change according to functional communication measure (FCM) of the ASHA National Outcome Measure System (NOMS) was equal to approximately 14 hours of direct intervention. For the purposes of this investigation, one level of FCM improvement would correspond to a treatment response for a single misarticulated phoneme such as /s/. Treatment studies that can yield a treatment response - defined here as greater than 70% accuracy in producing the /s/ sound in words and words-in-sentences (Van Riper & Emerick, 1984) in pre-treatment and post-treatment assessments – in a duration of therapy that is significantly less than the mean of 14 hours reported by Jacoby et al (2002), would suggest a comparative cost savings for public and private payers of speech therapy services.

In general, the more quickly the client is stimulable to the target behavior (in this case, remediating misarticulated /s/), the more efficiently can therapy proceed toward remediation (Bernthal, Bankson & Flipsen, 2009). One increasingly popular means of accelerating stimulability to achieve therapy gains is sensory biofeedback (McAllister Byun & Hitchcock, 2012). Sensory biofeedback in speech production utilizes specially designed instrumentation that facilitates increased awareness of the target behavior in the client. In addition, sensory biofeedback provides an external focus of directed attention to the task of remediating misarticulated speech. This external focus is said to aid the retention of a newly acquired motor skill such as speech (Wulf, 2007).

An example of sensory biofeedback utilizing primarily the client's visual sensory system (termed visual biofeedback) is ultrasound. In ultrasound for speech therapy, a transducer is placed under the chin which along with a linked software interface, displays a real-time image of the surface of the speech articulators inside the oral cavity. This image allows the client to more effectively contrast his own aberrant production with that of the clinician's model of correct production (Bernhardt, Gick, Bacsfalvi & Ashdown, 2003). Ultrasound has shown positive results in remediating residual, treatment-resistant /r/ errors (Adler-Bock, Bernhardt, Gick & Bacsfalvi, 2007), and in speech disorders associated with hearing impairment (Bernhardt et al, 2003). Other visual biofeedback approaches that have a strong base of evidence supporting their use are electromagnetic articulography (Katz,Bharadwai & Carstens, 1999; Wong, Murdoch & Whelan, 2010); electropalatography (Carter & Edwards, 2004; Lee, Law & Gibbon, 2009; McAuliffe & Cornwell, 2008); visual-spectral biofeedback (Shuster, Ruscello & Smith, 1992; Shuster, Ruscello & Toth, 1995; McAllister Byun & Hitchcock, 2012).

Recent studies have highlighted the strong connection between auditory and tactile or somatosensory feedback in speech perception and production (Tremblay, Shiller & Ostry, 2003; Gick & Derrick, 2009; Champoux, Shiller & Zatorre, 2011). For example, Gick and Derrick (2009) provided evidence that speech perception naturally includes a somatosensory component by demonstrating that inaudible tactile input during the perception of the voiceless bilabial stop interferes with normal adults' ability to perceive either /b/ or /p/ in a listening task. Tremblay, Shiller and Ostry (2003) found that just as humans use hearing to correct and refine speech production, they also use expected somatosensory patterns in everyday speech. In other words, if speech doesn't "feel right," then speakers will adjust their oral and facial movements to conform to how they expect speech to feel. As a natural corollary to these studies, leading psycholinguistic models of the speech production mechanism, such as Guenther's DIVA model (Guenther & Vladusich, 2012)

necessarily include a somatosensory feedback control subsystem that is "active during speech if the speaker's tactile and proprioceptive feedback from the vocal tract deviates from the somatosensory target region for the sound being produced" (p. 416). Moreover, given that consonant sounds such as /s/ require specific articulatory contacts within the vocal tract, fine-tuning a speaker's somatosensory acuity during speech production may be particularly important for consonant sounds such as /ʃ/ and /s/ (Ghosh, Matthies, Maas, Hanson, Tiede, Menard, Guenther, Lane & Perkell, 2010).

The treatment methodology of tactile biofeedback places a physical target within the oral cavity that enables a client to feel correct tongue placement and movement, and thereby refine his or her motor speech behaviors for the target sound. This additional, tactile information can then be used by the client and SLP to allow the client to perceive the correct placement of the speech articulators for a target sound and more efficiently achieve correct placement relatively early in the therapy process (Ruscello, 1995). The clinical promise of, specifically, tactile biofeedback in the treatment of articulation disorders has been the subject of previous investigations. One such investigation, Clark, Schwarz and Blakeley (1993), examined the efficacy of a tool embodiment of tactile feedback in the treatment of misarticulated American English /r/. The tool was fabricated in the form of a dental mold specially fitted with palatal targets for the tongue. This dental mold required individual fitting and fabrication for each study participant. Results of the study suggest that while the use of this tactile biofeedback embodiment was efficacious in the acquisition of treatment-resistant /r/, drawbacks were noted. Specifically, the dental mold design required costly individual fitting from an orthodontist or general dentist; the tool was reported by participants to be generally uncomfortable and, in some cases, was reported to have impeded saliva swallowing. The current study aims to provide further empirical validation for the methodology of tactile biofeedback by obtaining similar results to Clark, Schwarz, and Blakeley (1993), but with an optimized, better tolerated tool embodiment targeting the /s/ phoneme.

The focus of the current study is correcting misarticulated /s/ in neurotypical, hearing, native-English, pediatric speakers. Phonetically, /s/ is a voiceless sibilant fricative consonant requiring the speaker to effect a narrow constriction in the lingua-alveolar region of the oral cavity (Kent & Read, 2002) during production of this sound. If the precise placement of the tongue is not realized, the necessary oral constriction will in turn not be achieved and the resulting production of /s/ will sound distorted to the listener. Given this required precision, the /s/ phoneme is a commonly targeted error sound in speech therapy (Gibbon & Hardcastle, 1987; Bernthal, Bankson & Flipsen, 2009) and one that may require a variety of approaches to effectively remediate (Bleile, 2004). Ghosh et al (2010) provided evidence that somatosensory feedback may be especially valuable to the production of the English sibilant fricatives, /s/ and /ʃ/. Given these factors, /s/ was selected as the target error phoneme for this study.

The authors posit that the treatment methodology that most directly exploits the innate somatosensory speech control system described above is intra-oral tactile biofeedback. By incorporating a novel tool embodiment into therapy that consistently provides a precise lingual placement target within the oral cavity, study subjects will more effectively engage their innate somatosensory feedback mechanism in speech production and more efficiently remediate misarticulated /s/. The current study further aims to contribute to the comparatively modest research corpus focused on tactile biofeedback as well as to the treatment literature focused on short duration treatment regimens.

Hypothesis and Purpose

The purpose of this study is to report preliminary effectiveness data on the use of a specially designed intra-oral tactile biofeedback device targeting /s/. The study proposes to test the extent to which a group of children who exhibited phonetic-based sound system errors more efficiently remediated misarticulated /s/ given this consistently applied, precise lingual placement target within the mouth during treatment. To accomplish this, the authors implemented a randomized, controlled, single blind research protocol.

The study authors hypothesize: 1) the study group utilizing intra-oral tactile biofeedback in this short-duration therapy regimen will result in a treatment response, whereas the study group treated according to traditional methods of articulation therapy will not; and, 2) intra-oral tactile biofeedback therapy, delivered as a short-duration therapy regimen, will yield greater accuracy in production of /s/, as compared to traditional methods of articulation therapy.

METHOD

Figure 1. The test device and its components.

Test Device

The principal function of the intra-oral tactile biofeedback device, hereafter termed the test device, is to aid the participant in achieving correct lingual placement for the /s/ sound. As shown in Figure 1, the tip of the device suspends a small target 8 mm posterior to the front face of the upper front dentition. The target is a set of concentric circles 4 mm and 2 mm in diameter. While small, this target is easy to feel and provides tactile biofeedback to the user. These design parameters were optimized to most closely match correct lingual placement required across a variety of anatomical configurations. The device's dental stop and centering notch enable the consistent placement of the device target within the oral cavity, 8 mm posterior to the upper dentition. The test device prevents excessively anterior tongue placement (i.e. a frontal lisp or dental lisp) while cuing placement that is anterior enough in the oral cavity to yield a sound acoustically distinct from /ʃ/ (Bickford & Floyd, 2006). While there is some variation in the placement of the tongue tip to achieve a correctly perceived production of /s/ (McLeod, Roberts & Sita, 2006), a general tongue placement configuration is a helpful training tool in remediating /s/ (McAuliffe & Cornwell, 2008).

The device does not aid in achieving frication, or in preventing the phonological processes of stopping or initial voicing. It also does not help achieve lateral tongue bracing or medial lingual grooving, oral behaviors that are considered important components of achieving correct production of /s/ (McLeod, Roberts & Sita, 2006).

The target node for the tongue tip, and the shaft of device, were also designed to minimize airflow impedance during production of /s/. Since airflow parameters are critical to the production of /s/ (McLeod, Roberts, & Sita, 2006), careful consideration was paid to the device's ability not to impede airflow. In addition, the device was designed to minimally impede the coarticulation of other phonemes in a target item such as a word. These low impedance characteristics allowed participants to correctly produce /s/ up to the word level, with the test device in place.

The device is hand-held and was designed to be maximally controllable by the clinician while providing a direct tactile target for the placement of the tongue tip during production. To use the test device, the SLP places the tooth stop against the participant's upper front teeth, with the centering ridge used as an aid to place the device in the center of the participant's upper dentition. With the test device in place, the participant was instructed to place his or her tongue on the target to achieve correct placement for /s/. Figure 2 illustrates the use of the device once it is in place.

The tip of the device, which goes into the mouth, is made of a soft, thermoplastic elastomer that has passed appropriate biocompatibility and toxicity testing required by International Organization for Standardization (IOS) 10993 standards and the U. S. Food and Drug Administration (FDA). The material is soft enough to prevent deformation or pain when bitten down upon, yet is sturdy enough to retain its shape when manipulated by the tongue. The test device used in this study was the Speech Buddy® for /s/, designed by Articulate Technologies, Inc. (San Francisco, California, USA).

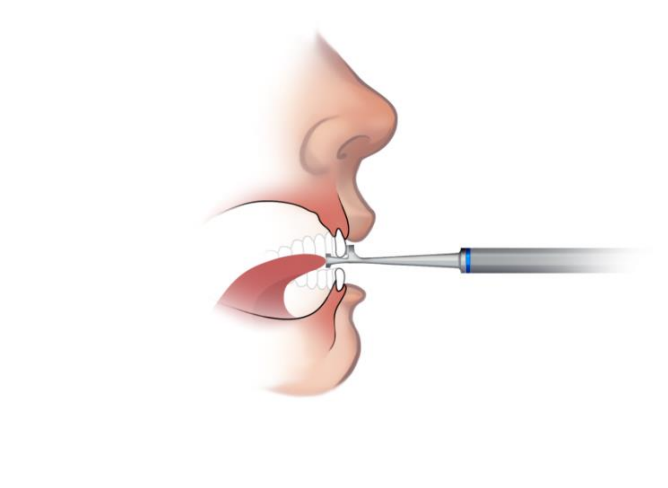

Figure 2. The placement of the test device within the oral cavity.

Participants

Twenty participants between the ages of 5:0 years and 8:11 years were enrolled in this study. Participants were recruited by advertisements in local print and online media, a directed mailer to local parents, and postings on local parenting internet listservs. All enrolled participants met the following criteria: 1) age 5:0 years 8:11 years at the time of assent and parental permission; 2) incorrect production of the /s/ phoneme (i.e. 0-20% correct) according to a picture naming test that contained 50 items; 3) receptive and expressive language skills greater than 1.5 standard deviations below the mean, as compared to normative values of the Comprehensive Evaluation of Language Fundamentals-4 Screening Test (CELF-4 Screening Test); 4) hearing function within normal limits at 500, 1000 and 2000 Hz in both ears, according to a hearing screening conducted in a quiet room with an Earscan 3 ® brand audiometer, calibrated at the beginning of the study; 5) native speakers of American English, according to pre-screening interviews with participants' parents and caregivers and the principal investigator's judgment of native language based on verbal interaction with the prospective participant; and, 6) have received less than ten hours of therapy time for a speech sound disorder that could be /s/ or any other sound, as per parent reports. The ten hours standard was specified in the protocol to aid in recruiting if necessary, but all 20 participants enrolled had no prior therapy addressing speech or language deficits.

Once screened, evaluated and enrolled in the study, each eligible participant was randomly assigned to a control group or an experimental group. There were two different speech language pathologist roles in the study. Four SLPs administered eight,

cost-free therapy sessions to the subjects, while one separate SLP, who was blinded as to which group the participant was in, performed accuracy assessments. The control group (n = 10) was administered traditional methods of articulation therapy during the eight sessions, whereas the experimental group (n = 10) was administered intra-oral tactile biofeedback as its primary cuing modality during the eight sessions. In addition to the therapy sessions, accuracy assessments were obtained before and after treatment, as well as at three specific intervals during the therapy regimen. As opposed to the therapy sessions, the accuracy of each participant's production of /s/ was determined by a different, single evaluator who was blind to the participant's inclusion in either the control or experimental group. Neither the control nor the experimental group used the test device during the accuracy assessments, which enabled to evaluator to be blinded.

All assessment and therapy sessions were conducted at Open Lines Speech and Communication, PLLC, a private practice facility in New York City. The study and site were fully approved by The Copernicus Group Institutional Review Board (IRB), the study's governing IRB. The study protocol, randomization protocol, data collection methods and statistical analysis, data storage, data forms, consenting procedures and potential conflicts of interest were all reviewed by the governing IRB. In addition, all recruitment materials, including ads in local media outlets and internet parenting listservs, were approved by the governing IRB.

Enrollment Summary and Randomization

A total of 24 children were identified and screened but only 20 met the inclusion criteria listed above. All pre-screenings were conducted by the principal investigator over the telephone once the participant's parent or caregiver made initial contact. All baseline evaluations to determine ultimate eligibility and to gather baseline data for the participants were conducted by a New York State-licensed, ASHA-certified Ph.D.-level clinician with over ten years of clinical experience who also acted as the study's single-blind evaluator.

Upon enrollment, participants were randomly assigned to the control or experimental group using the envelope method. Prior to the initiation of the study, 20 envelopes were created that included a visible sequence number (#01 to #20) on the outside of the envelope. The envelopes were opened in sequential order of randomization requests. A third party without knowledge of the study was identified to disclose randomization assignments. Randomization was not statistically driven and included include ten "experimental group" determinations and ten "control" group determinations to achieve a one-to-one randomization ratio.

Of the 20 participants enrolled in the intent-to-treat population, 15 participants were included in the per-protocol analysis due to the following reasons: loss to follow up, loss of upper front dentition during the study, concurrent therapy disclosed post-randomization. Participants are divided among the control group (n = 7) and the experimental group (n = 8). Table I summarizes participant characteristics.

Although study participants were required to have less than ten hours of previous speech therapy, all study participants had never received speech and/or language therapy in any form, targeting any deficit area. Prospective participants presenting with primarily phonologically based production errors were not excluded from participation; baseline accuracy of 0-20% was the only production-related inclusion criterion that determined. However, all enrolled participants presented with articulation-based production errors. There were no significant differences among the groups as to age, gender, raw score on the CELF-4 Screening Test (this test does not provide standard scores), pre-treatment accuracy percentage, type of /s/ distortion (frontal versus lateral), and elapsed time between the first and last therapy sessions, and amount of previous therapy, as shown in Table 1.

Description of the Eight Therapy Sessions

The goal of the therapy sessions was to administer nearly identical therapy to both the control and experimental group, with the exception being the fact that the experimental group used the test device, whereas the control group did not. The two groups received eight individual treatment sessions over a period of four to seven weeks. The study PI attempted to schedule two weekly sessions over four weeks. However, taking into account scheduling conflicts (e.g. vacation and illness), seven weeks was allotted to complete all eight sessions.

For both groups, each of the eight treatment sessions consisted of 45 total stimulus items, taking approximately 25 minutes to complete. The first five items trained auditory discrimination and asked the participant to discriminate between a correct versus incorrect production of /s/. The next six items trained were "warm-up" items with /s/ presented in isolation and in CV and VC syllables. After completing the "warm-up" items, the remaining 34 items trained /s/ in words in initial position (16 items), medial position (five items), and final position (13 items). These 34 stimulus items were used, randomly selected from a list of 140 total items (70 items in with /s/ in initial position, 15 items in medial position, 55 items in final position). Items were chosen to generally feature /s/ in stressed syllables and only as a singleton, and not in consonant clusters. Items were chosen to represent a wide range of vocalic and consonantal contexts. The total number of items trained was consistent for both test groups in all therapy sessions. Appendix A provides a sample therapy session, including randomized stimulus items.

The control group was treated according to a treatment manual which was developed according to phonetic-based practice principles stipulated in Van Riper (1978) and focusing primarily on phonetic placement cues. Each stimulus item began with phonetic placement techniques that described and visually illustrated to the participant correct placement. This was

Numerical Characteristics	Control Group			Experimental Group			Analysis		
	N	mean	SD	N	mean	SD	*t*	Df	*p*
Age at Baseline (months)	7	89.4	18.2	8	76.6	10.1	1.7	13	0.11
CELF-4 Screening Test [a]	7	9.0	2.9	8	7.0	3.1	1.3	13	0.22
Baseline Accuracy (%)	7	1.7	4.5	8	0	0	1.07	13	0.30
Time Between First and Last Therapy Session (days)	7	37.6	9.9	8	32.5	8.5	1.1	13	0.30
Prior Therapy (hours)	7	0	0	8	0	0	0.0	13	1

Binary Characteristics	Control Group			Experimental Group			Analysis
	N	N with characteristic	%	N	N with characteristic	%	*p*
Frontal Lisp	7	6	85.7	8	8	100.0	0.47 [b]
Male	7	3	42.9	8	3	37.5	1.00 [b]

[a] Number above criterion score.
[b] Fisher's Exact Test, two tailed.

Table 1. Characteristics of research participants.

followed by the clinician producing a model of the target sound in isolation. Sound discrimination was used to contrast the target sound with the participant's error production. After correct production of the target /s/ in isolation and discrimination of the target versus the error, practice was incorporated at the word level. In each session training was conducted at the isolation, syllable and word level according to the session description above.

Therapy in the experimental group proceeded according the treatment received by the control group, except that the primary cuing mechanism was intra-oral tactile biofeedback delivered by the test device. All "warm-up" items used the intra-oral tactile biofeedback device. In addition, every other item was trained with the intra-oral tactile biofeedback device, with 17 of the 34 total items trained with the device. Practice with the device at the isolation, syllable and word levels did not constrain movement of the speech articulators. As per the experimental group treatment manual, each child was provided with his or her own dedicated test device. After each production, the subject received immediate reinforcement from the study treating clinician on whether the production of /s/ in that item was correct or incorrect; and, if incorrect, what the primary reason for the misarticulation was (e.g. the tongue was misplaced between the teeth). Each clinician recorded whether a given item was correct or incorrect in the "therapy session log" document in each participant's trial binder. The recorded judgments of accuracy were *not* included in study assessments or data analyses reported below. After each session, clinicians were instructed to thoroughly wash the test device using mild soap and water.

Each participant's progress through the study was tracked by the principal investigator using dedicated trial binders consisting of all relevant study information for each participant. Each binder consisted of: executed parent and student consent forms; eligibility checklist to ensure fidelity with study inclusion and exclusion criteria; all pre-treatment, during-treatment and post-treatment assessments; all therapy session logs; protocol deviation reports; narrative summaries describing events that may have had a material impact on study data; adverse event reports; device malfunction reports, and post-trial questionnaires (experimental group only). The PI conducted periodic treatment fidelity checks which included weekly reviews of all trial binders as well as periodic in-session observations and telephone conversations with study clinicians.

Assessments/Measures

All accuracy assessment data were collected by a single ASHA-certified, licensed, Ph.D.-level evaluator with over ten years of clinical experience. The evaluator was blind as to the participant's inclusion in either the experimental or control group. The baseline and final assessments were 50 word assessments whereas the during-treatment assessments consisted of 20 words.

The assessment stimulus items were comprised of a picture-naming test containing pictures of objects or basic actions

containing /s/ in various words and in words-in-sentences, in various positions (initial, medial and final) and in various phonetic contexts. The same 50-word picture-naming test was used for the baseline assessment (the pre-treatment measure) and the final assessment (post-treatment measure). The during-treatment measures consisted of three separate 20-word tests consisting of randomly selected words from a set of 60 words. 50 of these words were the same 50 words comprising the baseline and final assessments, with an additional ten words included. No assessment items were ever used as treatment items. A list of assessment battery stimulus items can be found in Appendix B. Each participant was seen individually in a quiet therapy room for all assessments. During assessments, each participant was required to name each picture individually and each target response was scored by the single-blind assessor as either correct or incorrect, and then recorded on carbon paper.

To establish the reliability of the study evaluator, additional testing was conducted. Reliability testing used audio and video recordings of study stimulus items. The audio recorder used was the microphone attached to a JVC Everio GZ-MS120BU brand digital camcorder. The microphone had an audio sampling rate of 40 kHz, considered adequate for recording the entire acoustic signal of /s/ (Kent & Read, 2002). To establish intra-rater reliability, the evaluator was asked to judge these recorded study items on two separate occasions, 14 days apart. The same set of items was used, although they were presented in different random orderings. Inter-rater reliability was assessed using this method, which compared the accuracy judgments of the study evaluator to those of another judge, an experienced academic researcher, clinician, and former head of a leading accredited graduate clinic.

In addition, a qualitative questionnaire was given to experimental group participants at the conclusion of treatment. The questionnaire compiled participant impressions of their use of the test device. Questions posed to participants included: whether participants enjoyed using the device, whether they found the device scary, painful or uncomfortable, and if they felt the device helped them to speak better.

Clinician Training

The therapy sessions were conducted by four New York State-licensed and ASHA-certified speech-language pathologists, each with at least five years of clinical experience in treating speech sound disorders in a variety of clinical settings. The speech-language pathologists were trained on the relevant aspects of the study's protocol, on how to use the intra-oral tactile biofeedback device in therapy, how to conduct therapy sessions according to traditional methods of articulation therapy, and how to perform data collection for each study group. Each speech-language pathologist received a training manual so that the information could be consulted at a later date.

The dedicated training session consisted of a 30-minute presentation conducted by the study PI and was supported by detailed instructions with concise descriptions and visual supports. Instructions were provided for number of items, types of cuing permitted for each experimental condition, and type of reinforcement to be used. Clinicians were instructed on the recording of data as to correct versus incorrect production from the participant for each item on carbon paper, though these data were not included in any formal analysis. In addition, this training session included a segment devoted to the study therapists practicing to use the test devices with each other. Cleaning and storage protocols were also covered in the training session.

Statistical Analysis

Analyses were conducted using SAS (Statistical Analysis Software), Version 9.2 (SAS Institute Inc., Cary, NC, USA, www.sas.com). Independent sample t-tests and Chi-square analyses were conducted as appropriate, to test for differences between experimental and control group participants on the demographic and clinical characteristics shown in Table I. A repeated-measures Analysis of Covariance (ANCOVA) was then conducted to understand the effect of treatment with the intra-oral tactile biofeedback device on accuracy over time. Repeated-measures ANCOVA allows for both between-subject factors (in this case, treatment), and within-subject factors (in this study, time), to be tested. This statistical analysis allowed the authors to track improvements in performance among all participants with time, due to treatment, and to test treatment-specific effects, and differences in the trajectories of improvement over time, between control and experimental group members, and thus allowed the authors to more confidently attribute improvements in production accuracy of /s/ to treatment received.

RESULTS

The primary endpoint of the study, the determination of a response, or no response to therapy, was met ($p < .004$), indicating that, analyzed as a group, the participants using the test device experienced a treatment response, whereas the control participants, as a group, did not. The secondary endpoint, a statistical difference ($p < .05$) between the control and experimental group at a 95% confidence interval, was not met ($p = .08$). This indicates that the comparative improvement the experimental group experienced over the control group was statistically significant only to a 90% confidence interval, not the generally accepted standard of 95%.

Table 2 shows the raw data collected from the participants as well as the mean accuracy and standard deviation of each group at each time point. Figure 3 shows the graph of the mean accuracy at each time point.

As shown in Table 2 and Figure 3, the control and experimental groups showed similar baseline performance and response to untrained items at the first and second interim assessments; however, there is a change at subsequent probes. In addition, the

Subject	Baseline	Interim #1	Interim #2	Interim #3	Final	Change
			Experimental Group			
01	0 (0%)	2 (10%)	5 (25%)	16 (80%)	37 (74%)	(74%)
03	0 (0%)	10 (50%)	17 (85%)	19 (95%)	48 (96%)	(96%)
05	0 (0%)	7 (35%)	9 (45%)	16 (80%)	44 (88%)	(88%)
07	0 (0%)	10 (50%)	10 (50%)	12 (60%)	37 (74%)	(74%)
09	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	(0%)
11	0 (0%)	0 (0%)	0 (0%)	17 (85%)	37 (74%)	(74%)
14	0 (0%)	16 (80%)	N/A	18 (90%)	44 (88%)	(88%)
19	0 (0%)	16 (80%)	20 (100%)	20 (100%)	49 (98%)	(98%)
Mean % (SD)	0.0 % (0.0)	38.1% (32.7)	43.6% (38.9)	73.8% (32.2)	74.0% (31.5)	74.0% (31.5)
			Control Group			
02	0 (0%)	8 (40%)	10 (50%)	10 (50%)	14 (28%)	(28%)
06	0 (0%)	15 (75%)	20 (100%)	20 (100%)	50 (100%)	(100%)
12	0 (0%)	0 (0%)	0 (0%)	1 (5%)	0 (0%)	(0%)
15	0 (0%)	0 (90%)	0 (0%)	0 (0%)	0 (0%)	(0%)
16	6 (12%)	N/A	N/A	11 (55%)	45 (90%)	(78%)
17	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	(0%)
18	0 (0%)	20 (100%)	20 (100%)	20 (100%)	50 (100%)	(100%)
Mean % (SD)	1.7% (4.5)	35.8% (42.1)	41.7% (47.6)	44.3% (44.4)	45.4% (49.1)	43.7% (47.4)

Table 2. Data set of assessment accuracy: number correct (percentage correct).

	n	Response	No Response
Experimental Group	8	7 (87.5%)	1 (12.5%)
Control Group	7	3 (42.8%)	4 (57.1%)

Table 3. Response using threshold of 70% accuracy.

experimental group demonstrated increased response accuracy, while the control group plateaued. The control group showed minimal improvement in probe testing after the second interim assessment. However, the experimental group showed a continued increase in mean accuracy after the first half of treatment, as measured by the third assessment and the final assessment. The change in accuracy from the baseline to the final assessment was greater for the experimental group (mean = 74.0%, S.D. = 31.5, n = 8) than for the control group (mean = 43.7%, S.D 47.4, n =7). Variance in treatment response within the control group was considerable due to the inconsistent and binary pattern of response rate; three subjects showed a measurable treatment response, whereas four showed no measurable treatment response.

Statistical Analysis

The primary hypothesis was tested using a one-way repeated-measures ANCOVA to compare performance over time for the experimental group versus the control group to determine the effect of the intra-oral tactile biofeedback therapy. Performance at baseline was the covariate. There was a significant interaction between time and group, $F(3,35) = 5.46$, $p = .004$, showing that the experimental group demonstrated a significant response, while the control group did not.

Post hoc comparisons between different group-time conditions showed that the experimental group performed significantly better at the final assessment and third interim assessment than at the first interim assessment ($t = 3.99$, $p < .001$, $t = 3.97$, $p < .001$, respectively). Additionally, the experimental group also performed significantly better at final assessment and third interim assessment than at the second interim assessment (t =

2.96, p = .006, t = 2.94, p = .006, respectively). No other post-hoc comparisons between different time and group combinations were statistically significant.

A two sample t-test was used to examine the secondary hypothesis of the difference between the control and experimental group. A one sided t-test was used and assumed equal variances between the groups. At $\alpha = .05$ the change was not statistically significant ($p > .05$) with $t = 1.48$ and $p = .08$, while a significant response was generated using $\alpha = .10$, indicating a trend finding.

In addition to the per-protocol analysis, an additional analysis was performed on the intent-to-treat population of n = 20, using the method of last measured observation carried forward for the additional five participants using the same statistical methods outlined above. The results also yielded a similar, insignificant response for $\alpha = .05$ with $t = 1.52$ and $p = .07$.

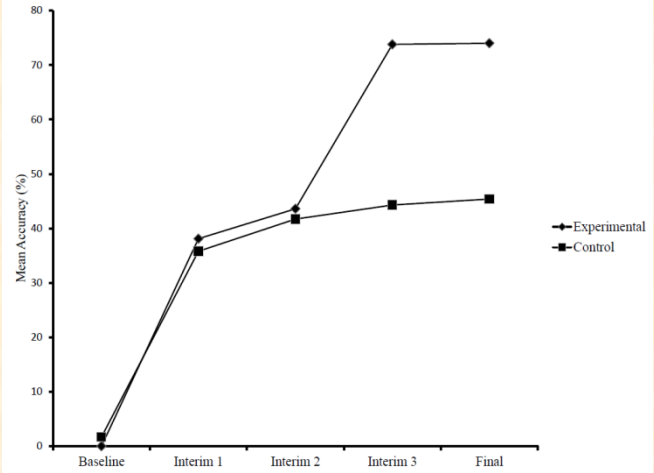

Figure 3. Mean accuracy of production of /s/ at each time point.

Rater Reliability Results

Intra-rater correspondence among the recorded study stimulus items was 95%. Inter-rater correspondence among the items was 94%. According to the standards in the field set by McCauley and Swisher (1984), these reliability measures were in the acceptable range and suggest the study blinded evaluator was a reliable judge of correct versus incorrect /s/ production.

Post-Trial Questionnaire Results

The post-trial questionnaire showed that all of the participants had no adverse reactions to the device, there were no device malfunctions, and no participants found the tool uncomfortable or painful. One participant reported not to have enjoyed using the device.

DISCUSSION

The results of this investigation suggest that the addition of an intra-oral tactile biofeedback device enabled significant gains across the short-duration treatment period and support the test device's use as a viable treatment option in articulation therapy. Results of qualitative participant questionnaires also showed that none of the experimental group participants reported that the device was painful, uncomfortable or scary. This further suggests that the current embodiment of the intra-oral tactile biofeedback device has the potential to be a generally well-tolerated clinical tool by neurotypical children. In addition, all study subjects were naïve to treatment at enrollment. Generally, the inclusion of treatment-naïve subjects in behavioral research eliminates the effect of prior treatment on observed results (He, Deng, Li, Chen, Jiang, Wang, Huang, Collier, Gong, Ma, Zhang & Li, 2012). This would lend support to the assertion that the changes observed in subjects' accuracy of /s/ production in this study were directly due to the therapy they received.

The results reported above show that participants can be identified as responders and non-responders. Table 3 shows responders and non-responders, using a threshold of 70% production accuracy on the pre-treatment vs. post-treatment assessments, based from a 70-80% performance criteria accuracy range identified by Van Riper and Emerick (1984). The consistency of the experimental group's treatment response provides further evidence of the utility of biofeedback approaches in treating articulation disorders. Seven out of eight participants in the experimental group were identified as responders, and one participant with 0% accuracy at final assessment was deemed a non-responder. Conversely, in the control group, only three out of seven participants were identified as responders with the remaining four showing little to no response. There were no baseline or otherwise measured clinical characteristics that could distinguish the non-responders in either group. This rate of non-response of 43% grossly corresponds to the 28% rate of non-response observed by Jacoby et al (2002).

For the three responders in the control group, traditional methods of articulation therapy may have been all these children needed to learn and habituate correct /s/. For the other four non-responders, either traditional methods were not adequate to achieve a learning breakthrough, or more time in therapy was required. In either case, that seven of eight experimental group participants responded to treatment in this short-duration therapy regimen would suggest that having access to all potentially useful sensory components for learning (i.e. including the tactile component) may increase the chances that a short-duration therapy program would be effective.

Despite their consistent misarticulations of /s/, at the outset of treatment all study participants were able to achieve the necessary sibilant frication for /s/ and no participant manifested any commonly described phonological process (e.g. stopping). Therefore, all participants' attempted productions of /s/ were realized as incorrect lingual placement within the oral cavity. For both experimental conditions, according to each group's treatment manual, the focus of therapy was on training correct

oral configuration, and differentiating participants' trained, correct realizations of /s/ versus previous, incorrect realizations of /s/. The treatment gains achieved by control group participants can be attributed to the effectiveness of traditional methods in achieving clinical gains. Given that the experimental group's primary cuing method was intra-oral tactile biofeedback, the treatment gains achieved by the experimental group can be attributed to the ability of the device to position the tongue tip in the correct region of the oral cavity while simultaneously not impeding the airflow necessary to achieve frication for /s/. Verbal directions were used to enhance the salience of the new, correct oral configuration for /s/. In addition, cues were used to aid the participants in auditorily discriminating correct and incorrect acoustic realizations of /s/. While these techniques are hallmarks of traditional methods of articulation therapy (Van Riper & Emerick, 1984), the primary learning modality in the experimental group was tactile biofeedback delivered via the test device.

The results reported above also provide preliminary evidence supporting the therapeutic corollary to the somatosensory feedback mechanism described in experiments involving normal control adults not presenting with speech sound deficits (Champoux, Shiller & Zatorre, 2011; Gick & Derrick, 2009; Tremblay, Shiller & Ostry, 2003). In addition, these results provide further clinical support of the validity of a distinct somatosensory input in psycholinguistic models of speech production, such as Guenther's DIVA model. The results also seem to corroborate suggestions by Ghosh et al (2010) that given the precision required to produce sibilant fricatives, this class of speech sounds may be especially appropriate targets for therapy involving somatosensory feedback.

Despite these apparent strengths and the significant findings with respect to the study's first hypothesis, the inclusion of additional subjects may have resulted in a significant (e.g. $p < .05$) finding for the study's secondary hypothesis. The study was limited by the cost of study personnel and the inclusion of treatment-naïve participants, which significantly extended the study's recruitment calendar. Absent these constraints, the enrollment of an additional eight to ten participants may have made the results more definitive.

As per the description of study participants above, this study included only pediatric participants whose language and hearing functions were determined to be within normal limits. The data obtained would apply to those children who present with a similar cognitive, language and hearing profile. Additional studies are required to determine whether these results can be applied to those with hearing impairments, concomitant language disorders or cognitive impairments.

This study was designed to examine the effect of intra-oral tactile biofeedback in treating solely misarticulated /s/. No other speech errors in participants' speech sound systems were systematically assessed or treated. For this reason, a standardized assessment that may have revealed severity of involvement and level of stimulability for the production of /s/ was not included in the study's assessment battery. While the majority of participants presented with /s/ as their sole misarticulated phoneme, the lack of attention to the participants' other potential treatment needs would suggest the results obtained may not be generalized to treatment outcomes targeting the child's whole speech sound system.

Traditional methods of articulation therapy were selected as the treatment control for this study. This was due to traditional methods' comparative advantage in achieving favorable treatment outcomes in treating /s/ (Powell, Elbert, Miccio, Strike-Roussos & Brasseur, 1998). In addition, the authors felt that the use of the test device for /s/ was most complementary with traditional methods, rather than, for example, phonological approaches. In contrast, other studies have found that phonological therapy approaches are comparatively effective (e.g. Pamplona, Ysunza & Espinosa, 1999). The results obtained in this study would not bear upon this apparent discrepancy and are best interpreted as preliminary evidence supporting tactile biofeedback in itself, rather than evidence supporting tactile biofeedback in lieu of a particular leading treatment methodology.

This is a preliminary study designed to examine the clinical utility of an intra-oral tactile biofeedback device in remediating misarticulation of the /s/ phoneme in a cohort of neurotypical, hearing children. The results reported above were achieved in a short duration therapy period and suggest that intra-oral tactile biofeedback has the potential to reduce the expected time required to treat misarticulated /s/. In closing, it is worth noting that the majority of pediatric speech sound disorders may be effectively and efficiently treated by traditional methods of articulation therapy (Powell et al, 1998). However, a significant number of pediatric subjects remain resistant to treatment despite the use of these traditional approaches (e.g. 28% of pre-school and school-age children, according to Jacoby et al, 2002). The results obtained in this study provide evidence for intra-oral tactile biofeedback as a cost-effective alternative to traditional approaches to articulation therapy, when such approaches have previously failed to achieve desired clinical results.

ACKNOWLEDGMENTS

The authors wish to thank: Dr. Dennis Ruscello, for his mentorship and for reviewing drafts of the manuscript; Dr. Tara McAllister Byun for reviewing drafts of the manuscript; and, Dr. John Bernthal for his guidance in determining appropriate study design. The authors would also like to thank the study clinicians, Sandi Chmielewski, M.S. CCC-SLP, Carol Ferrone, Ph.D. CCC-SLP, Renee Fisher, M.S., CCC-SLP, and Melanie Yukov, M.S. CCC-SLP, and to Genise Burnett, research assistant to the first author. The study was sponsored by Articulate Technologies, Inc. The Copernicus Group IRB oversaw the ethical handling of participants and quality of data collection and analysis.

CORRESPONDENCE

gordy@articulatetech.com
mschesin@gmail.com

DECLARATION OF INTEREST

The Copernicus Group IRB monitored the study to ensure that there were no conflicts of interest that would jeopardize subject enrollment, data collection, and data analysis. The first author is co-founder and Chief Scientific Officer of Articulate Technologies, Inc. and is compensated on a part-time, consultant, fee-for-service basis, and by an equity stake in the company. The statistical analysis was performed by the second author, an independent consultant, who does not have a financial interest in the study sponsor.

REFERENCES

Almost, D., & Rosenbaum, P. (2000). Effectiveness of intervention for phonological disorders: a randomized controlled trial. *Developmental Medicine & Child Neurology*, 40, 319-325.

American Speech-Language-Hearing Association (2008). *Evidence-Based practice* (EBP). website: www.asha.org/members/ebp

Bacsfalvi, P., & Bernhardt B. M. (2011). Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: ultrasound and electropalatography. *Clinical Linguistics & Phonetics*, 25, 1034-43.

Bernhardt, B., Gick, B., Bacsfalvi, P., & Adler-Bock, M. (2005). Ultrasound in speech therapy with adolescents and adults. *Clinical Linguistics and Phonetics*, 19, 605-617.

Bernhardt, B., Gick, B. Bacsfalvi, P. & Ashdown, J. (2003). Speech habilitation of hard of hearing adolescents using electropalatography and ultrasound as evaluated by trained listeners. *Clinical Linguistics and Phonetics*. 17, 199-217.

Bernthal, J. E., Bankson, N. W., & Flipsen, P. (2009). Articulation and Phonological Disorders: Speech Sound Disorders in Children. Boston, MA: Pearson.

Bickford, A.C. & Floyd, R. (2006), Articulatory Phonetics: Tools for Analyzing the World's Languages. Austin, TX: SIL International.

Bleile, K. M. (2004). Manual of Articulation and Phonological Disorders: Infancy through Adulthood. Clifton Park, NY: Thomson/Delmar Learning.

Broen, P. A., & Westman, M. J. (1990). Project parent: a preschool speech program implemented by parents. *Journal of Speech and Hearing Disorders*, 55, 495-502.

Byun, T. M., & Hitchcock, E. R. (2012). Investigating the use of traditional and spectral biofeedback approaches to intervention for /e/ misarticulation. *American Journal of Speech-Language Pathology*, 21, 207-221.

Carter, P., & Edwards, S., (2004). EPG therapy for children with long-standing speech disorders: predictions and outcomes. *Clinical Linguistics and Phonetics*, 18, 359-372.

Champoux, F., Shiller, D. F., & Zatorre, R. J. (2011). Feel what you say: an auditory effect on somatosensory perception. PLoS ONE 6(8). e22829. doi:10.1371/journal.pone.0022829.

Clark, C. E., Schwarz, I. E., & Blakeley, R.W. (1993). The removable R-appliance as a practice device to facilitate correct production of /r/. *American Journal of Speech-Language Pathology*, 2, 84-92.

Crowe Hall, B. J. (1991). Attitudes of fourth and sixth graders toward peers with mild articulation disorders. *Language Speech and Hearing Services in Schools*, 22, 334-340.

Dickson, K., Marshall, M., Boyle, J., McCartney, E., O'Hare, A., & Forbes, J. (2009). Cost analysis of direct versus indirect and individual versus group modes of manual-based speech-and-language therapy for primary school-age children with primary language impairment. *International Journal of Communication Disorders*, 44, 369-381.

Eiserman, W. D., McCoun, M., & Escobar, C. M. (1990). A cost-effectiveness analysis to two alternative program models for serving speech-disordered preschoolers. *Journal of Early Intervention*, 14, 297-317.

Ghosh, S. S., Matthies, M. L., Maas, E., Hanson, A., Tiede, M., Menard, L., Guenther, F. H., Lane, H., & Perkell, J. S. (2010). An investigation of the relation between sibilant production and somatosensory auditory acuity. *Journal of the Acoustic Society of America*, 5, 3079-3087.

Gibbon, F.E. & Hardcastle, W.J. (1987). Articulatory description and treatment of 'lateral /s/' using electropalatography: a case study, *British Journal of Disorders of Communication*, 22, 203-217.

Gick, B. & Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*, 462, 502-504.

Gierut J.A. (1998). Natural domains of cyclicity in phonological acquisition. *Clinical Linguistics and Phonetics*, 12, 481-499.

Guenther, F. H., & Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of Neurolinguistics*, 25, 408-422.

He, Z., Deng, M., Li, Z., Chen, L., Jiang, Q., Wang, C., Huang, D., Collier, A., Gong, Q., Ma, N., Zhang, T., & Li, T. (2012). Aberrant intrinsic brain activity and cognitive deficit in first episode treatment-naïve patients with schizophrenia, *Psychological Medicine*, 10, 1-12.

Jacoby, G., Lee, L., Kummer, A., Levin, K., & Creaghead, N. (2002). The number of individual treatment units necessary to facilitate functional communication to improvements in the speech and language of young children. *American Journal of Speech-Language Pathology*, 11, 370-380.

Katz, W. F., Bharadwaj, S. V., & Carstens, B. (1999). Electromagnetic articulography treatment for an adult

with Broca's Aphasia and Apraxia of Speech, *Journal of Speech, Language and Hearing Research*, 42, 1355-1366.

Kent, R. D. (1997). The Speech Sciences. San Diego, CA: Singular Publishing Group, Inc.

Kent, R. D., & Read, C. (2002). Acoustic analysis of speech. San Diego, CA: Singular Thomsen Learning.

Klein, E. S. (1996). Phonological/Traditional approaches to articulation therapy: a retrospective group comparison. *Language, Speech and Hearing Services in Schools*, 27, 413-323.

Lee, A. S., Law, J., & Gibbon, F. E. (2009). Electropalatography for articulation disorders associated with cleft palate. *Cochrane Database of Systematic Reviews*, 8, CD006854.

Major, E. M., & Bernhardt, B. H. (1998). Metaphonological skills of children with phonological disorders before and after phonological and metaphonological intervention. *International Journal of Language & Communication Disorders*, 1998, 413-444.

McCauley, R. J. & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49,* 34-42.

McLeod, S., Roberts, A., & Sita, J. (2006). Tongue/palate contact for the production of /s/ and /z/. *Clinical Linguistics & Phonetics*, 51-66.

Nasir, S., & Ostry, B. (2006). Somatosensory precision in speech production. *Current Biology*, 16,1918-1923.

Overby, M., Carrell, T., & Bernthal, J. (2007). Teachers' perception of students with speech sound disorders: a quantitative and qualitative analysis. *Language Speech and Hearing Services in Schools,* 38, 327-341.

Pamplona, M. C., Ysunza, A., & Espinosa, J. (1999). A comparative trial of two modalities of speech intervention for compensatory articulation in cleft palate children, phonologic approach versus articulatory approach. *International Journal of Pediatric Otolaryngology*, 49, 21-26.

Powell, T. W., Elbert, M., Miccio, A. W., Strike-Roussos, C., & Brasseur, J. (1998). Facilitating [s] production in young children: an experimental evaluation of motoric and conceptual treatment approaches. *Clinical Linguistics & Phonetics*, 12, 127-146.

Rubens, R.J. (2000), Redefining the survival of the fittest: communication disorders in the 21st Century. *Laryngoscope*, 110, 241-245.

Ruscello, D. M. (1995). Speech appliances in the treatment of phonological disorders. *Journal of Communication Disorders*, 28, 331-353.

Rvachew, S., & Nowak, M. (2001). The effect of target-selection strategy of phonological learning. *Journal of Speech, Language, and Hearing Research*, 44, 610-623.

Semel, E., Wiig, E. H., & Secord, W. A. (2004). Clinical evaluation of language fundamentals, fourth edition—Screening test *(CELF-4 screening test).* Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Shriberg, L.D. & Kwiatkowski, *J. (1994).* Developmental phonological disorders I: clinical profile. *Journal of Speech and Hearing Research, 37, 1100-1126.*

Shuster, L. I., Ruscello, D. M., & Toth, A. R. (1995). The use of visual feedback to elicit correct /r/. *American Journal of Speech-Language Pathology*, 4, 37-45.

Tremblay, S., Shiller, D. M., & Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, 866-869.

Tyler, A. A., & Lewis, K. E. (2005). Relationships among consistency/variability and other phonological measures over time. *Topics in Language Disorders*, 25, 243-253.

Van Riper, C. (1978). Speech Correction: Principles and Methods. Englewood Cliffs, NJ: Prentice-Hall.

Van Riper, C., & Emerick, L. (1984). Speech Correction: An Introduction to Speech Pathology and Audiology. Englewood Cliffs, NJ: Prentice-Hall, 1984

Wong, M. N., Murdoch, B. E., & Whelan, B. (2010). Kinematic analysis of lingual function in dysarthric speakers with Parkinson's disease: an electromagnetic articulograph study. *Journal of Medical Speech-Language Pathology*, 12, 414-425.

Wulf, G. (2007), Attentional focus and motor learning: A review of 10 years of research. *E-journal Bewegung und Training*, 1, 4-11.

AUTHOR INFORMATION

Gordy Rogers, M.S. CCC-SLP holds a B.A. in linguistics (Yale University) and an M.S. in speech-language pathology (Teachers College Columbia University). He is clinical founder and Chief Scientific Officer of Articulate Technologies, Inc., the developer of the Speech Buddies system for treating speech sound disorders. In addition, Mr. Rogers is owner of Village Speech Solutions, PLLC, a private practice focusing on speech sound disorders and based in New York City. Mr. Rogers is a frequent contributor to industry-related conferences and publications.

Megan Chesin is a Postdoctoral Research Fellow at Columbia University/New York State Psychiatric Institute, Department of Psychiatry, Molecular Imaging and Neuropathology Division. Her research interests include modeling and testing pathways to psychopathology and adapting and testing treatments for depressed and other clinical populations. She employs path analyses and generalized linear mixed models.

APPENDIX A
Sample Session and Guide to Use of Stimulus Cue

Cue Number	Cue	Type	Control Group	Experimental Group
		Auditory Discrimination		
1	sip - tip	n/a	traditional cue	traditional cue
2	thin - sin	n/a	traditional cue	traditional cue
3	sack - Zack	n/a	traditional cue	traditional cue
4	walrus - walruh	n/a	traditional cue	traditional cue
5	fussy - futhy	n/a	traditional cue	traditional cue
		Warm Up		
6	S	isolation	traditional cue	tactile cue
7	S	isolation	traditional cue	tactile cue
8	suh	initial syllables	traditional cue	tactile cue
9	suh	initial syllables	traditional cue	tactile cue
10	us	final syllables	traditional cue	tactile cue
11	us	final syllables	traditional cue	tactile cue
		Therapy		
12	south	initial	traditional cue	tactile cue
13	saga	initial	traditional cue	traditional cue
14	cinnamon	initial	traditional cue	tactile cue
15	self	initial	traditional cue	traditional cue
16	syrup	initial	traditional cue	tactile cue
17	sickle	initial	traditional cue	traditional cue
18	send	initial	traditional cue	tactile cue
19	sat	initial	traditional cue	traditional cue
20	sap	initial	traditional cue	tactile cue
21	city	initial	traditional cue	traditional cue
22	safe	initial	traditional cue	tactile cue
23	soil	initial	traditional cue	traditional cue
24	silver	initial	traditional cue	tactile cue
25	simple	initial	traditional cue	traditional cue
26	city	initial	traditional cue	tactile cue
27	said	initial	traditional cue	traditional cue
28	fossil	medial	traditional cue	tactile cue
29	lesson	medial	traditional cue	traditional cue
30	juicy	medial	traditional cue	tactile cue
31	recipe	medial	traditional cue	traditional cue
32	wrestle	medial	traditional cue	tactile cue
33	princess	final	traditional cue	traditional cue
34	grace	final	traditional cue	tactile cue
35	chase	final	traditional cue	traditional cue
36	loss	final	traditional cue	tactile cue
37	mass	final	traditional cue	traditional cue
38	gross	final	traditional cue	tactile cue
39	grease	final	traditional cue	traditional cue
40	class	final	traditional cue	tactile cue
41	across	final	traditional cue	traditional cue
42	hiss	final	traditional cue	tactile cue
43	chase	final	traditional cue	traditional cue
44	brace	final	traditional cue	tactile cue
45	ace	Final	traditional cue	traditional cue

APPENDIX B
50 Word Baseline and Final Assessments

Cue Number	Cue	Word position
/s/ in Words		
1	Sock	initial
2	Sun	initial
3	Six	initial
4	Seal	initial
5	Cereal	initial
6	Salt	initial
7	Saw	initial
8	Seven	initial
9	Sing	initial
10	Circle	initial
11	Soup	initial
12	Sink	initial
13	Submarine	initial
14	Sick	initial
15	Soccer Ball	initial
16	Castle	medial
17	Messy	medial
18	Muscle	medial
19	Motorcycle	medial
20	Dinosaur	medial
21	Missile	medial
22	Glasses	medial
23	Sausages	medial
24	Eraser	medial
25	Medicine	medial
26	Dress	final
27	Mice	final
28	Chess	final
29	Glass	final
30	Rice	final
31	(Shoe) Lace	final
32	Ice	final
33	Grass	final
34	Goose	final
35	House	final
36	Octopus	final
37	Cactus	final
38	Gas	final
39	Face	final
40	Bus	final
/s/ in Words in Sentences		
41	Sit(ting)	initial
42	Santa (Claus)	initial
43	Sandwich	initial
44	Soap	initial
45	Whistle	medial
46	Baseball	medial
47	Mouse	final
48	Dice	final
49	Lettuce	final
50	Moss	final

EFFECTS OF ANXIETY ON VOICE PRODUCTION: A RETROSPECTIVE CASE REPORT OF SELECTIVE MUTISM

Cesar E. Ruiz
Evelyn R. Klein
LaSalle University

ABSTRACT

Selective mutism (SM) is a disorder that presents in childhood whereby children exhibit a failure to speak in specific situations despite the ability to speak when feeling comfortable. This perplexing disorder is thought to be related to anxiety. The purpose of this study was to report on the experiences of a young adult male who wanted to speak but failed to do so because of his SM during his childhood and adolescence. In particular, the authors focused on his recount of vocal symptoms and strategies used to compensate for his difficulty communicating. Data from the Selective Mutism Questionnaire, comprehensive voice analysis, and past history in addition to his personal reflections and how he overcame the disorder are presented.

KEY WORDS

Selective Mutism
Voice
Social Anxiety

BACKGROUND

According to the 2013 Diagnostic and Statistical Manual of Mental Disorders (DSM-5), the diagnosis of selective mutism (SM) is given to children from early childhood to adolescence who meet the following criteria: (a) An inability to speak in at least one specific social situation where speaking is expected (e.g., at school) despite speaking in other situations (e.g., at home); (b) The disturbance has interfered with educational or occupational achievement or with social communication; (c) The duration of the selective mutism is at least one month and is not limited to the first month of school; (d) The failure to speak is not due to a lack of knowledge of or comfort with the primary language required in the social situation; and, (e) The disturbance cannot better be accounted for by a communication disorder (e.g. childhood onset fluency disorder) and does not occur exclusively during the course of a pervasive developmental disorder, schizophrenia or other psychotic disorder (APA, 2013).

Epidemiological studies estimate that the age of SM onset ranges from 2.7 years to 4.1 years (Garcia, Freeman, Miller, & Leonard, 2004; Giddan, Ross, Sechler, & Becker, 1997). This childhood disorder affects approximately 1 in 140 children from various cultures around the world with an average 2 to 1 ratio of girls to boys with SM (Viana, Beidel, & Rabian, 2009). Children who are timid and anxious tend to be more susceptible (Black & Uhde, 1995). Interestingly, there is no single identified cause of SM although social anxiety has been implicated and manifestations of SM can vary with some children remaining mute outside the home while others talk to select individuals in specific situations (Black, 1996; Chavira, Shipon-Blum, Hitchcock, Cohan, and Stein, 2007; Cunningham, McHolm, and Boyle, 2006; Cunningham, McHolm, Boyle, and Patel, 2004).

Anxiety disorders have been reported to affect between 5% and 10% of children and adolescents and are considered to be one of the most prevalent disorders for this age group (Davis, May, & Whiting, 2011). Selective mutism has been associated with anxiety disorders and described as an extreme form of social phobia (Black, 1996; Black & Uhde, 1995; Kristensen, 2000a; Kristensen, 2000b, Steinhausen, Wachter, Laimböck, & Metzke, 2006). Not all investigators agree. Omdal & Galloway (2008) questioned the description of SM as a social phobia. They argued that if that were the case, children should respond to Cognitive Behavioral Therapy (CBT) much better than they currently do. Based on their interviews of six adults who recovered from SM as children, the authors suggested that the term *specific phobia of expressive speech* may be more descriptive of the reported behaviors. The children interviewed did not report behaviors of isolation or withdrawal in social situations but rather of avoidance, especially when they knew that the situation would require them to speak.

While anxiety is a major feature of SM, there are other vulnerabilities that make an individual more susceptible to the disorder. Family immigration status, bilingualism, and communication delays have been reported as contributors. Elizur and Perednik (2003) reported a higher incident of SM among immigrants (2.2%) when compared to the general population (0.76%). The authors further indicated that the increase in prevalence was associated with a combination of social anxiety and environmental vulnerability. Toppelberg, Tabors, Coggins, Lum, and Burger (2005) indicated that SM is at least three times higher in bilingual children. They listed risk-aversive behaviors, inhibited temperamental disposition, and language delays as possible contributing factors. They cautioned however that SM should not be confused with the normal silent period that children may go through when exposed to a second language. They added that this period may extend up to 6 months. Communication delays have also been reported as a contributing factor. Klein, Armstrong, and Shipon-Blum (2012) identified an underlying expressive language deficit during narratives. The authors noted that the SM children's verbal scores decreased as language became more complex.

In an attempt to fill the gap between what has been reported about children with SM, and the children's own perspectives, Omdal (2007) interviewed six adults who had suffered from SM as children. The participants' responses were categorized into the following themes: (a) The origin of SM involved all participants reporting that they withdrew from social interactions, especially if they found the situations to be traumatic; (b) SM was able to be maintained as being mute became part of their manner of socialization; (c) Children gained a sense of control over their environment by not talking and to talk would bring about change and produce discomfort; and (d) Staying mute was safe but became non-functional as demands and expectations changed. All six individuals interviewed reported making the decision to consciously break the silence and thereby created a safe environment to begin the process. The final theme dealt with psychological adjustment; (e) Despite having overcome their selective mutism, the participants reported the need for psychological intervention to help them adjust to the trauma the whole experience of SM caused.

Much research on selective mutism has focused on qualification and quantification of the children's behaviors in unfamiliar situations (I e., the school) compared to familiar situations (I e., in the home) (Kristensen, 2000b; Steinhausen, Watcher, Laimbock, and Metzke, 2006; Yaganeh, Beidel, Turnar, Pina, and Silverman, 2003) as communication is in sharp contrast in those settings. In addition, investigators have focused on data from treatment specialists, parents engaged in interviews, and nonverbal observational interactions with the children (Cunningham, McHolm, Boyle, and Patel, 2004; Dummit, Klein, Tancer, Asche, & Martin, 1997; Manassis, Fung, Tannock, Sloman, Fiksenbaum, & McInnes, 2003; Manassis, Tannock, Garland, Minde, McInnes, & Clark, 2007). Little research exists from children's perspectives, in part due to their communicative difficulties. However, having the opportunity to speak to a young adult who was selectively mute gave us the ability to use

a retrospective case study design (Street & Ward, 2010) . This type of research included three factors: (1) data was collected after the events occurred, (2) there was access to both first-person accounts and data, and (3) the final outcomes were known when the data collection took place. Using this approach, the purpose of our study was to add to the literature about SM from an historical account of factors that impacted an individual's life on his way to recovery.

METHOD

Participant

The participant in this case review was a 30-year old white male, currently working full-time as a market research analyst. He grew up in a middle-class home with a mother who stayed home, a father who worked full-time as a business manager, and a sister two years older who grew up to become a social worker. Home life had its challenges with a father who had obsessive-compulsive personality and parents who argued in front of their children. At school Lenny was bright and learned with ease. He performed well in class with the exception of being mute.

At school Lenny seldom spoke to a peer and didn't speak when called on by his teacher. He said he never spoke to a group of peers and didn't ask his teacher any questions. With his family he often spoke comfortably at home but he seldom spoke to his family when in unfamiliar places or when others were nearby. He also refrained from talking to relatives that did not live with him. He seldom spoke in situations where there were organized teams or clubs outside of school. Not talking in social situations greatly interfered with his life and caused him extreme feelings of discomfort.

Lenny recalled himself as being a child who was very self-conscious and perfectionistic. He reported himself as a child who was ordinary in appearance, wearing glasses from the age of 10 years old. On the *Burns Perfectionism Scale* (Burns, 1980) Lenny remembered that he set very high standards for himself and believed he should be upset if he made a mistake. He also indicated that it was shameful to display weakness or foolish behavior. He thought that failing at something important meant he was less of a person. He also thought that if he scolded himself for failing to live up to his expectations that it would help him to do better in the future.

From Lenny's experiences, he said that he responded better to adults who showed some amount of vulnerability; meaning they were okay to expose themselves to judgment regarding social norms. For example, they may have disclosed their vulnerability through small actions such as acting silly, pretending to be overwhelmed in a comical sense and generally demonstrating to him that they were not perfect which meant that he did not have to be perfect either. Committing a social faux pas was helpful in that Lenny saw the more human side of the person instead of an adult who could be threatening.

When others focused on Lenny he felt more anxious. Talking would have brought too much attention and potential vulnerability. In retrospect he stated that there was nothing at stake to bring him to use his voice in public. It was easier and acceptable to remain silent. Being acknowledged in public was very troubling to him. He recalled a time when he was six years old and struggled to take off his Halloween costume at school. When his teacher had to assist him in front of the class, he cried. The way he made it through school was to shut-down and stay alone. Verbal expression became physically difficult and full of anxiety that would go to his throat. Any sounds he attempted to utter came out garbled and saying words and sentences became nearly impossible. Lenny's muteness helped him avoid making mistakes or calling attention to himself. He felt that speaking was a threatening entity. Given all his anxiety and feelings of defeat, Lenny only spoke regularly at home with his parents and sister.

Procedures

The participant, Lenny, reached out to the board of directors of a national SM organization to share his personal experiences in hopes of helping others with SM who suffered as he did. At the time of the consult, Lenny had been speaking for many years and was ready to tell his story. After a phone interview it was suggested that he travel from New York to the university clinic in Philadelphia for a formal meeting and voice assessment which took place over a period of three hours with the authors. His description of what happened to his 'throat' when he tried to speak was perplexing.

The Interview. The semi-structured interview was conducted by the authors at the university clinic. Lenny was eager to share information from his past and present experiences. Information from his childhood revealed that he had been frequently bullied at school for acting 'girly' which further made him withdraw from talking. His parents ended up doing most of the talking for him. He believed it was easier to not talk and after a brief time he was no longer expected to talk. Lenny stated that he was 'enabled' not to talk. It was much easier to remain silent at school and in social settings. Always watching and always quiet, Lenny became accepted as the "quiet one." He felt that school was oppressive and in order to be around other children, he did his best to become unaware of them. Although teachers were concerned with Lenny's lack of speech, collaboration between home and school did not occur and he did not get professional help.

Voice Evaluation. Adequate voice production requires good coordination between the respiratory, phonatory, resonatory, and articulatory systems. That is, after taking a breath, the air is exhaled to activate vocal folds vibration, and the sound produced travels outwards either through the oral or nasal cavity to be used for speech purposes. This is a process that begins with our first cry at birth and continues as part of our existence as we grow older. Scharfstein, Beidel, Sims, & Finnell (2011) reported the presence of distinct vocal patterns among

children with social phobias. Their study revealed that children with social phobia tended to have a softer speaking voice and higher pitch. In our interview with Lenny, he reported that when he had to speak, he felt like "his throat" would close. He indicated that he knew exactly what he wanted to say but could not speak due to his perceived inability to get his voice started. It is speculated that the vocal folds may have become temporarily and excessively adducted (approximated in order to vibrate and produce sound) thereby preventing the necessary airflow to initiate vibration. His lack of understanding about the vocal mechanism accompanied by a feeling that the voice would not work in a particular setting (or with specific individuals) may have further increased anxiety and mutism.

Given the authors' clinical knowledge about the vocal mechanism, an attempt was made to obtain more information about Lenny's past and current vocal symptoms. A comprehensive perceptual and acoustic voice evaluation was performed. Lenny was asked to describe in more details what he felt as he attempted to produce voice during his years with SM and now. The *Multidimensional Voice Program* (MDVP) (KayPentax, 2011a) which measures vocal parameters during sustained phonation of /a/, and the *Real-Time Pitch* (KayPentax, 2011b) which also measures vocal parameters in continuous speech, were utilized to measure Lenny's vocal parameters as compared to normal ranges for his age and gender. Attempts were made to use surface electromyography (sEMG) to measure muscle tension; however, due to his beard, adequate measurements could not be obtained.

Lenny was later asked to try to imitate the vocal behaviors he experienced during his years with SM in an attempt to observe his reported vocal symptoms. The *Rainbow Passage* (Fairbanks, 1960) was utilized to facilitate verbal output. Digital manipulation of the larynx was performed to assess for tension sites and trigger points. An electrolarynx was also used to attempt to elicit speech without voicing. (An electrolarynx is a speech device that generates sounds when placed on the neck of an individual who then mouths the words without vibration of the vocal cords).

RESULTS

Diagnostic Confirmation

Using the *Selective Mutism Questionnaire* (SMQ) (Bergman, Keller, Piacentini, & Bergman, 2008), recollections of Lenny's elementary years regarding speaking can be found in Table 1 as they relate to his speaking experiences at school, with family, and in social situations outside of school. Lenny spoke 44% of the time when with his family. His most difficult situations involved social encounters outside school where he spoke only 7% of the time. In school he spoke 11% of the time compared with 88% of the time for age-matched peers. The SMQ is typically completed by the child's parents. In this situation information was obtained directly from Lenny, as the most reliable source.

Place	Raw Score	% Talking	% Talking Norms
At School	2 of possible 18	11%	88%
With Family	8 of possible 18	44%	97%
In Social Situations (outside school)	1 of possible 15	7%	83%
Total	11 of possible 51	21%	90%

Table 1. Elementary school recollections about speaking using the SMQ.

Vocal History

Lenny indicated that during speech tasks as a child he felt a sudden pressure in "his throat". When asked to point to the area he was describing, he grabbed each side of his larynx with his thumb and index finger and pressed inwards. He reported an increase of resistance to initiate his voice accompanied by a sensation of shortness of breath. He stated that over the years he learned to compensate for the reported pressure by producing the voiceless velar phoneme /k/. By repeating this phoneme several times, he felt he could maintain his airflow more consistent and voicing became easier. He was observed during this behavior and it was noted that the exaggerated breathiness indeed allowed for maintenance of glottic opening, but also it lowered the larynx decreasing laryngeal excursion during voicing.

Lenny readily agreed to imitate the vocal behaviors he experienced during his years with SM. As a result, his demeanor changed and a sense of anxiety was observed. He stared at the investigators with a look of fear as if predicting an uncomfortable experience. Once he began reading the *Rainbow Passage* he exhibited hesitations. Lenny visibly struggled as he attempted to verbalize fluently. As Lenny attempted to speak, his larynx was noted to abnormally elevate from the lower resting position. Assessment of tension sites and trigger points elicited pain and discomfort to pressure in the thyrohyoid space bilaterally, especially posteriorly. The electrolarynx was tried to determine if Lenny could use the device to elicit speech without vocal tension. The investigator held the device on the right side of his larynx and activated it so Lenny could speak as he read the *Rainbow Passage* without using his voice. Lenny exhibited difficulty disengaging his voice and using the one produced by the electrolarynx. His anxiety built up to the point that he uttered "take it away" despite minimal pressure on the larynx, and minimal soft vibration from the device.

Voice Analysis

Results from the *Multidimensional Voice Program* (KayPentax (2011a). were as follows: Average Fundamental Frequency (F_o)= 107 Hz. (normal = 100-135 Hz.); Jitter= 0.3% (normal <1%) and Shimmer= 2.1% (normal <3.8%). The *Real-Time Pitch Program* KayPentax (2011b). elicited the following results: F_o= 99 Hz.; Pitch Range= 553 Hz. (high) and 93 Hz. (low); Maximum Phonation Time of /a/ was 9 seconds (normal 15-20 seconds); s/z ratio of 1.2 (normal 1 - 1.2) with /s/= 16.65 seconds and /z/= 14.45 seconds; and Mean Loudness of 63.5 dB (ranged from 50 – 71 dB) with a standard deviation = 4.54. At the time of testing, all voice parameters were judged to be within normal limits.

A Specialized All-Body Approach

Lenny indicated that in his search for help to overcome his condition, he came across a vocal technique known as *The Linklater*. He reported that as he began training using this technique, his focus on voicing shifted from his throat to his whole body and that imaging facilitated his vocal control. *The Linklater Voice Method* (Linklater, 2006) incorporates a training technique of freeing the natural voice. It was originally developed by Kristin Linklater, a professor in the Theater Department at Columbia University, to help actors who needed to project their voices in film. The training has been referred to as emotionally rigorous but rewarding. The technique incorporates body movements as well as use of nasal passages and vocal cords. According to Dru Pilmer (2013), a professor at Indiana University Southeast, the Linklater Voice Method helps the person breathe down into the belly. It helps people to use a natural voice that has a two to four octave pitch range.

It is known that vibrations of voice carry information and that as humans we modify sounds to communicate. It is interesting to think about infants who even without words can let others know their message. It is the physical tension that is said to inhibit the message carried by voice. That is precisely what Linklater works on conveying. Her method has been said to help the individual re-learn natural breathing to reduce vocal tension and explore the free range of sound making. Breath capacity, lung volume, and speech articulation are all emphasized. There is an awareness of the whole body.

In Linklater's book, *Freeing the Natural Voice* (2006), it is said that when the voice is free, the person is free. Interrupted breathing patterns are reduced which help reduce stress and tension. Exertion doesn't seem to affect the trained voice and people with such training seem to be able to captivate attention via their voice. They are also said to listen better.

In 2012, Emily Wessendarp, a graduate student in speech-language pathology from the University of Toledo, wrote a thesis on the effects of the *Linklater Voice Training Technique* on the voices of student actors. Although the study only had eight subjects (four who received the Linklater Training and four matched acting students who were in a no treatment control group), those in the experimental Linklater group demonstrated significant benefits regarding maximum phonation time compared to the control group. Those in the Linklater group exhibited improved airflow control during sustained phonation. Voice projection was positively impacted. However, Wessendarp noted that those in the experimental group, showed a trend for shimmer which may require further monitoring.

DISCUSSION

Reflections

Lenny reflected on his experience as a person with SM. His willingness to share his own story was prompted by his hope to educate others and help alleviate their pain. He offered four pieces of advice.

First, he suggested that parents foster their children's interest in activities and hobbies by participating with them one-on-one. During that time he encouraged parents try to promote the knowledge that it is okay to make mistakes because that is how people learn.

The second suggestion was for parents to find play dates for their children. This may be a peer who is a family friend or a relative who is sensitive and kind and who would adapt their behavior to complement the child with SM. Lenny suggested that the sensitive friend be someone from outside the school setting during the initial stages of verbal communication. Having an outside school friend could give the child with SM an outlet and a peer with whom to play without the worry of having to duplicate talking in the school environment. Lenny believed that his outside school friend helped him alleviate some stress and gave him something to look forward to prior to making a friend in school. Unfortunately, he said his friend lived too far away and he seldom had the chance to see him. He recalled feeling very lonely.

Third, Lenny suggested that parents (as well as teachers) boost the SM child's confidence by acknowledging a real accomplishment (as opposed to an overall personal quality statement) and to do this without requiring the child to speak. While praising the child for a specific act is good, Lenny cautioned against asking the child to tell someone about it. He believed that would have caused him extreme anxiety as children with SM generally want to avoid personal attention.

Fourth, Lenny strongly believed in the need to alleviate physiological tension. As a child with SM, he said he was usually tense. Although he reported the primary tension at his mid anterior neck, he believed that whole body relaxation techniques were helpful. Lenny stated that using the *Linklater Voice Technique* helped him open up physiologically which continued to help him emotionally. Gaining voluntary vocal control, which is for most of us an involuntary task, proved effective through the technique. As he moved his focus from his larynx to his whole body, voicing became part of the total picture rather than only a source for speech. (As authors, we are not promoting this

technique but rather report on it from the client's perspective). Lenny revealed that up to that point he thought he couldn't talk when in fact it was hard but not impossible. He said that having more knowledge about the vocal mechanism for the purpose of speech would have been helpful. Such a reflection may serve as an explanation as to why it may be so difficult for some children with SM to describe their vocal symptoms. In addition, it may also provide another direction for evaluators to consider during their assessments.

Throughout Lenny's reflection, he repeatedly expressed the concept of enabling and referred to it as one of the biggest problems in maintaining his SM. Due to his accompanying anxiety, it was more comfortable for him not to speak. Parents and teachers filled in the gaps for his muteness and nonverbal communication and it became a way of life for him throughout school and in most social settings. He began to be known as "the child who didn't talk." Lenny said he didn't realize the importance of talking and since talking would have further heightened his anxiety, he remained silent. It wasn't until he was employed and about to lose his job that he knew he had to talk. Losing that job would have meant financial instability. His fear of being unemployed outweighed his fear of speaking. With the support of several months on a selective serotonin reuptake inhibitor (SSRI) medication and voice therapy, Lenny began to use his voice. He has since remained employed and has gone on to be a fully functioning adult speaker.

Conclusion

The information gathered during this interview proved to be valuable in detailing the emotional, physiological and psychological aspects of one individual's experiences with SM. Hearing the description from an educated man who had the experience of living with SM brings to light the complexity of this disorder. In summary, several themes emerged: (a) speech was avoided rather than social encounters being avoided as would be the case with social phobia. Lenny reported not being afraid of the social situation but of the possibility of having to speak in that situation; (b) enabling by parents and teachers made it possible for Lenny not to talk (which he preferred as that helped diminish his anxiety); (c) he had difficulty initiating his voice despite knowing what he wanted to say; and (d) treatment ultimately helped him overcome his fear of speaking.

The common denominator for these themes appeared to be the effect that anxiety had on Lenny's ability to initiate voice. If heightened anxiety can increase heart rate, reduce saliva, and increase perspiration, it may be possible that for some individuals, it can also increase laryngeal tension and compromise vocal fold vibration. Once Lenny stopped focusing on voice production at the laryngeal level and recognized that respiration and resonation were also part of voicing, he began to gain control of his voice initiation and thus his consistent ability to produce speech. He believed that the natural breathing treatment approach helped him reduce vocal tension. He also found that emphasizing increased breath capacity, resonance, and speech articulation helped him more readily initiate speech sound production. He gained confidence to produce consistent and a controlled voice. As Lenny gained control in his ability to vocalize, he no longer avoided situations.

Based on this case study and information about anxiety and its potential effects on voicing, it is recommended that instrumental (laryngeal surface electromyography (sEMG), videostroboscopy, flexible laryngoscopy, acoustic voice analysis, etc.) and/or non-instrumental assessment of vocal parameters (s/z ratio, breath support for phonation, maximum phonation time, vocal quality, etc.) be incorporated into the evaluation. The possibility of vocal tension interfering with voice production due to anxiety is new and requires further investigation for assessment and treatment of SM.

CORRESPONDENCE

Cesar E. Ruiz
ruiz@lasalle.edu
Evelyn R. Klein
Klein@lasalle.edu

REFERENCES

American Psychiatric Association (2013). Diagnostic and statistical manual disorders *Fifth Edition.* Washington, DC: Author.

Bergman, R. L., Keller, M. L., Piacentini, J., & Bergman, A. J. (2008). The development and psychometric properties of the Selective Mutism Questionnaire. *Journal of Clinical Child and Adolescent Psychology, 37*(2), 456-464.

Black, B. (1996). Social anxiety and selective mutism. *American Psychiatric Press Review of Psychiatry, 15*, 469-495.

Black, B. & Uhde, T.W. (Jul 1995). Psychiatric characteristics of children with selective mutism: A pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry, 34(7),* 847-856.

Burns, D. (1980). The perfectionist's script for self-defeat. *Psychology Today*, November, 34-52.

Chavira, D. A., Shipon-Blum, E., Hitchcock, C., Cohan, S., & Stein, M. B. (2007). Selective mutism and social anxiety disorder: All in the family? *Journal of the American Academy of Child and Adolescent Psychiatry, 46,* 1464–1472.

Cunningham, C. E., McHolm, A., Boyle, M. H. and Patel, S. (2004). Behavioral and emotional adjustment, family functioning, academic performance, and social relationships in children with selective mutism. *Journal of Child Psychology and Psychiatry, 45*, 1363–1372. doi: 10.1111/j.1469-7610.2004.00327.x

Cunningham, C. E., McHolm, A. E., & Boyle, M. H. (2006). Social phobia, anxiety, oppositional behavior, social skills, and self-concept in children with specific

selective mutism, generalized selective mutism, and community controls. *European Child & Adolescent Psychiatry, 15,* 245–255.

Davis, T.E., May, A., & Whiting, S.E. (2011). Evidence-based treatment of anxiety and phobia in children and adolescents: Current status and effects on the emotional response. *Clinical Psychology Review, 31,* 592-602.

Dummit, E.S., Klein, R.G., Tancer, N.K., Asche, B., & Martin, J. (1997). Systematic assessment of 50 children with Selective Mutism. *Journal of the American Academy of Child & Adolescent Psychiatry, 36(5),* 653-660.

Elizur, Y., & Perednik, R. (2003). Prevalence and description of selective mutism in immigrant and native families: A controlled study. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(12), 1451-1459.

Fairbanks, G. (1960). The Rainbow Passage. *Voice and Articulation Drillbook (2nd Ed).* New York. NY. Harper & Row, Publishers. p. 127.

Garcia, A.M., Freeman, J.B., Miller, L.M., & Leonard, H.L. (2004). Selective Mutism (T.H. Olledick & J.S. March, Eds.). In Phobic and Anxiety Disorders in Children and Adolescents: A Clinician's Guide to Effective Psychosocial and Pharmacological Interventions (pp. 433-458).

Giddan, J. J., Ross, G. J., Sechler, L. L., & Becker, B. R. (1997). Selective mutism in elementary school: Multidisciplinary interventions. *Language, Speech, and Hearing Services in Schools, 28,* 127-133.

KayPentax (2011a). Multi-Dimensional Voice Program (MDVP), Model 5105. Retrieved on April 8, 2013 from http://www.kaypentax.com/index.php?option=com_product&controller =product &Itemid =3&cid[]=56&task=pro_details

KayPentax (2011b). Real-Time Pitch, Model 5121, Version 3.3. Retrieved on April 8, 2013 from http://www.kaypentax.com/index.php?option=com_product&controller=product&Itemid=3&cid[]=96&task=pro_details

Klein, E.R., Armstrong, S.L., & Shipon-Blum, E. (2012). Assessing spoken language competence in children with selective mutism: Using parents as test presenters. *Communication Disorders Quarterly, 20,* 1-12.

Kristensen, H. (2000a). Selective mutism and comorbidity with developmental disorder/delay, anxiety disorder, and elimination disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 39(2),* 249-256.

Kristensen, H. (2000b). Multiple informants report of emotional and behavioural problems in a nation-wide sample of selective mute children and controls. *European Child & Adolescent Psychiatry*, 10(2). 135-142.

Linklater, K. (2006). *Freeing the Natural Voice: Imagery and Art in the Practice of Voice and Language.* London, England: Nick Hern Books Limited.

Manassis, K., Tannock, R., Garland, E. J., Minde, K., McInnes, A., & Clark, S. (2007). The sounds of silence: Language, cognition and anxiety in selective mutism. *Journal of the American Academy of Child & Adolescent Psychiatry, 46,* 1187–1195.

Manassis, K., Fung, D., Tannock, R., Sloman, L, Fiksenbaum, L., & McInnes, A. (2003). Characterizing selective mutism: Is it more than social anxiety? *Depression & Anxiety*, 18(3), 153-161.

Omdal, H., & Galloway, D. (2008). Could selective mutism be re-conceptualized as a specific phobia of expressive speech? An exploratory post-hoc study. *Child and Adolescent Mental Health, 13(2),* 74-81.

Omdal, H. (2007). Can adults who have recovered from selective mutism in childhood and adolescence tell us anything about the nature of the condition and/or recovery from it? *European Journal of Special Needs Education, 22(3),* 237-253.

Pilmer, D. (2013, March 11). A completely different approach. *IU News Room.* Retrieved from http://newsinfo.iu.edu/web/page/normal/3320.html

Roeser, Pearson, & Tobey (1998). Rainbow Passage. *Speech-Language Pathology Desk Reference.* New York, NY: Thieme Publishers, p.362.

Scharfstein, L.A., Beidel, D.C., Sims, V.K., & Finnell, L.R. (2011). Social skills deficits and vocal characteristics of children with social phobias or Asperger's disorders: A comprehensive study. *Journal of Abnormal Child Psychology, 39,* 865-875.

Steinhausen, H.-C., Wachter, M., Laimböck, K. and Metzke, C. W. (2006), A long-term outcome study of selective mutism in childhood. *Journal of Child Psychology and Psychiatry, 47*, 751–756. doi: 10.1111/j.1469-7610.2005.01560.x

Street, C., & Ward, K. (2010). Retrospective case study. In A.J. Mills, G. Durepos, & E. Wiebe, (Eds.), *Encyclopedia of Case Study Research*. Sage Publications, Inc. DOI:http://dx.doi.org/10.4135/9781412957397.

Toppelberg, C.O., Tabors, P., Coggins, A., Lum, K., & Burger, C. (2005). Differential diagnosis of selective mutism in bilingual children. *Journal of The American Academy of Child and Adolescent Psychiatry, 44,* 592-595.

Viana, A.G., Beidel, D.C., & Rabian, B.(2009). Selective mutism: A review and integration of the last 15 years. *Clinical Psychology Review, 29,* 57-67.

Wessendarp, E. (2012). *Effects of the Linklater Voice Training Technique on the Voices of Student Actors*. Graduate thesis. Department of Speech-Language Pathology, University of Toledo, OH

Yeganeh, R., Beidel, D. C., Turner, S. M., Pina, A. A., & Silverman, W. K. (2003). Clinical distinctions between selective mutism and social phobia: An investigation of childhood psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 1069–1075.

AUTHOR INFORMATION

Cesar Ruiz, SLP.D. CCC-SLP, BRS-S, is an Associate Professor at LaSalle University, Philadelphia, PA. He is also a Board Recognized Specialist in Swallowing and Swallowing Disorders. Dr. Ruiz's clinical appointments include the Philadelphia VA Hospital and the Ear-Nose-Throat Association of Chester County, PA (ENTACC). He is an adjunct doctoral professor and dissertation chair at Nova Southeastern University in Fort Lauderdale, FL. His experience includes speech and swallowing rehabilitation following head and neck cancer, as well as service to other dysphagia and voice disorder patients. Dr. Ruiz is a trained nasendoscopist and provides workshops around the country to train others. He has published and presented nationally on a variety of topics in medical speech-language pathology.

Evelyn Klein, Ph.D. CCC-SLP, BRS-CL, is Professor of Speech-Language-Hearing Science at LaSalle University, Philadelphia, PA, and a licensed, certified speech-language pathologist and licensed psychologist with post-doctoral training in neuropsychology. She holds Board Specialty Recognition in Child Language and became an ASHA Fellow in 2009. In 2011, Dr. Klein received the Lindback Distinguished Teaching Award from LaSalle University, and in 2012 she was awarded the Pennsylvania Clinical Achievement Award from PSHA> Recent publications include *Focus on Function, Focus on Transition, Hear It-Say It-Learn It,* and the textbook *Acquired Language Disorders: A Case-Based Approach.* Dr. Klein is currently principal investigator of a grand-funded research program investigating speech and language skills in children with selective mutism.

TRAINING SPEECH-LANGUAGE PATHOLOGISTS FOR THE ASSESSMENT AND INTERVENTION FOR BILINGUAL CLIENTS

Koren Beardshall
Clarion University
Erin Lundblom
The University of Pittsburgh

ABSTRACT

The number of people who represent culturally and linguistically diverse backgrounds continues to increase in the United States. These individuals are often either over-assessed or under-assessed due to speech and language concerns. This may occur when speech-language pathologists (SLPs) are unfamiliar with how to best provide services. The purpose of this research study was to determine the extent to which SLPs believe graduate school prepares them to work with linguistically diverse clients. A 20-item questionnaire was developed and distributed to 395 SLPs in Pennsylvania employed in a school setting. The survey items were developed to collect information pertaining to the SLPs' experiences assessing and treating bilingual clients for speech and language, as well as, their educational background related to bilingual clients. The results of this study implicate the importance of graduate education and continued training for SLPs in Pennsylvania to meet the needs of linguistically diverse clients.

KEY WORDS

Cultural and Linguistic Diversity
Bilingual Assessment
Bilingual Treatment
Survey Research
Graduate Training

INTRODUCTION

Speech-language pathology is a field which continues to grow not only in the quantity of practitioners but also in its scope of practice. The Bureau of Labor Statistics predicts approximately 28,800 new openings between the years 2010 and 2020 (BLS, 2013). The field grows not only in the quantity of practitioners but also in its scope of practice, which is "driven by educational reform, legal mandates and evolving professional practices" that impacts the roles and responsibilities of SLPs in school-based settings (ASHA, 2010b). The number of people who speak multiple languages or come from culturally, diverse backgrounds residing in the United States continues to increase. Although the United States does not track bilingual individuals, the census asks several language related questions that can help estimate bilingualism (Grosjean, 2012). Approximately 51 million people in the United States speak a language other than English in the home, but also know English; this number represents roughly 18% of the U.S. population and does not include young children (Grosjean, 2012). Such an increase in bilingualism is coupled with a need for increased bilingual speech-language pathology services.

Currently the growing need for educators and related service professionals who can meet the needs of such diverse language users has not been realized. There is an ever-increasing shortage of teachers and other professionals, including speech-language pathologists (SLPs), who have the appropriate skills and training to meet the needs of bilingual clients (Annett, 2002). One estimate from the Occupational Outlook Handbook indicated that the need for speech-language pathologists with bilingual experience or expertise will increase 11% each year and result in an additional 30,000 jobs by 2016 (Bureau of Labor Statistics, 2008). According to the Pennsylvania Department of Education, in Pennsylvania there are 42,542 students in the school setting identified, as Limited English Proficiency (LEP) (Pennsylvania Department of Education, 2013).

With the increase in the number of linguistically and culturally diverse groups (U.S. Census Bureau, 2003), more and more often SLPs are asked to assess or identify a speech and language disorder for a child in a language other than the SLP's native language. In fact, many SLPs working in a school setting admit to not being proficient in providing services to culturally and linguistically diverse populations (ASHA, 2011). As a result, students who are bilingual are often either over-assessed or under-assessed for speech and language problems (Peña & Bedore, 2011). Bilingual students may be under assessed as having speech and/or language impairments because the SLP may believe that they are struggling with learning a second language rather than the problem representing an impairment (Limbos & Geva, 2001). In other instances, school professionals fail to take into account the problems which naturally occur in bilingual language development, and they place students in special education settings inappropriately. The student may be following the rules of their native language, and as, Geisler (2010) stated, "the majority of educators and school-based speech-language pathologists are accordingly unfamiliar with the process young children go through in learning a second language" (Geisler, 2010).

ASHA has offered guidance through numerous policy documents and statements regarding the provision of services appropriate for culturally and linguistically diverse populations (ASHA, 2004; ASHA, 2004b, ASHA 2010c, ASHA 2012). For example, ASHA (2004) outlined culturally and linguistically appropriate competencies that SLP's must meet to provide appropriate services. These standards include competencies in cultural competence (e.g. Influence of one's own beliefs and biases in providing services), language competencies of the clinician (e.g. Developing appropriate collaborative relationships with translators/interpreters), and specific disorder areas (e.g. language, articulation and phonology, resonance/voice/fluency, swallowing, and hearing/balance). For example, this information relates to the disorder area of language:

> *Appropriate use of published test materials in language assessment including standardized norm-referenced tests…including analyzing normative sampling limitations…issues related to validity and reliability and inherent cultural and linguistic biases in these test materials. Application of appropriate criteria so that assessment materials/tests/tools that fail to meet standards be used as informal probes, with no accompanying scores.*

A study by Caesar and Kohler (2007) explored how often standardized assessments versus authentic and flexible assessment procedures were used by SLPs with bilingual populations. The authors reported that SLPs "relied more on formal, standardized measures than on informal, alternative procedures" when testing bilingual students (Caesar & Kohler, 2007, p. 196). They concluded, "appropriate, nonbiased assessment of bilingual students, who are suspected of speech language difficulties has become one of the major challenges confronting school-based speech pathologists" (Caesar & Kohler, 2007, p. 190). Standardized, norm-referenced English tests have been demonstrated as inadequate for the assessment of bilingual students (Caesar & Kohler, 2007).

One alternative to standardized testing, which has been used regularly, is language sampling. Language sampling can assist clinicians when deciding between a language disorder and a language difference, in their linguistically diverse clients (Gutierrez-Clellen, Restrepo, Bedore, Pena, & Anderson 2000). In language sampling, a clinician analyses a conversational sample of speech obtained from the child. Technological advances in the past few years have made it easier as well as more efficient to use language sampling.

Another method used in the assessment of bilingual clients is dynamic assessment. Dynamic assessment typically involves a

test, teach, and then re-test approach. First the client's current level of performance is determined, then the SLP teaches the client strategies and practices the task, and then the client is again assessed and these results are compared to the pretest to determine progress (ASHA, 2012). A dynamic assessment approach differs from a standardized test in that it accounts for language and cognitive abilities as well as behavioral and motivational factors (Tzuriel, 2001). In an experimental study done by Kapantzoglou, and colleagues (2012), dynamic assessment was found to have "promising results for use as an alternative method for evaluating vocabulary learning skills in predominately Spanish speaking preschoolers" (p. 92).

ASHA has repeatedly supported alternative approaches including language sampling and dynamic assessment for culturally and linguistically diverse populations in a number of policy documents (ASHA, 2004; ASHA, 2004b, ASHA 2010c, ASHA 2012). As Caesar and Kohler (2007) reported, however, the two "practices used least frequently" are those which "apply directly to the assessment of bilingual children"—those practices include dynamic assessment and language sampling (Caesar and Kohler, 2007 p. 197). SLPs, specifically those in school-based settings, may not know how to use these assessments, or may not have the resources or experience to do so (Caesar and Kohler, 2007). A clinicians' experience may affect how he/she believes the assessment should occur when working with bilingual clients. Kritikos, (2003) explained, "experience may affect an SLP's beliefs about how to interpret and gather assessment data" (Kritikos, 2003). Dollaghan and Horner (2011) further explained that because there are no ideal measures to identify bilingual clients, "clinicians can justify using most of these measures in efforts to identify," however, "until more, stronger, and more precise evidence is available" the results should only be viewed as "suggestive of diagnostic status" (Dollaghan & Horner, 2011, p. 1086).

Given the multitude of assessment approaches and confounding problems, the issues SLPs have related to assessment may translate to an under and over assessment of bilingual clients. The ASHA scope of practice states that SLPs "are committed to the provision of culturally and linguistically appropriate services" and that one aspect of this is to determine whether the communication problem arises from a disorder, or from difficulty with learning a new language (ASHA, 2007). A language difference is, "a dialectal variety of American English" and should never be diagnosed as a disorder (ASHA, 2003). The task of differential diagnosis, however, may be difficult for speech-language pathologists who have limited training or experience in the assessment and treatment of bilingual clients.

This phenomenon of over- and under- assessing bilingual clients may be related to the lack of experience speech pathologists report in this area (ASHA, 2008). Researchers have demonstrated that SLPs are often not confident in their ability to work with multicultural clients. A survey completed by ASHA reported, "Only about 8% of school clinicians report training in bilingual assessment" (ASHA, 2008). The ASHA 2010 Schools Survey indicated that 8% of clinicians reported that they were 'not at all' qualified to work with multicultural populations, and 7.6% said they were 'very' qualified to work with this population (ASHA, 2010). The ASHA 2011 Membership Survey reported that 5% of SLPs were 'not at all' qualified to work with multicultural populations whereas 8.4% of SLPs believed they were 'very' qualified to work with the multicultural population (ASHA, 2011). Overall, consistency appears in these ratings. It is hypothesized that the presumed lack of confidence in assessing and treating bilingual clients may be related to graduate training programs. For example, bilingual speech-language assessment and treatment may not be sufficiently targeted at the preservice level. Therefore, many SLPs may not be able to recognize whether or not the student is showing typical signs of difficulty learning a second language or a language disorder.

The purpose of this study was to investigate to what extent, SLPs in Pennsylvania report: (a) training related to the assessment and treatment of bilingual clients and, (b) assessment and treatment experiences with bilingual clients.

METHODS

Design

A cross-sectional survey design was selected in an attempt to gather data about current practices. The design was used to study a group of speech-language pathologists working in Pennsylvania school-based settings. A questionnaire was created to learn more about the relation between the education and training practicing SLPs have received related to bilingual assessment and treatment, as well as their experiences with bilingual assessment and treatment during employment. The questionnaire contained both quantitative and qualitative questions to provide a better understanding of the data from the study.

Participants

Speech-language pathologists working in a school-based setting in the Pennsylvania who are members of the American Speech Language and Hearing Association (ASHA), a national professional organization for speech-language pathologists, audiologists, and teachers of the hearing impaired, were recruited to participate. A list of 300 names and accompanying physical mailing address were purchased through ASHA filtered to include only certified SLPs in the state of Pennsylvania who indicated a school-based setting as their primary employment location. An additional 95 names and addresses were obtained due to their affiliation with Clarion University of Pennsylvania as an externship supervisor in a school-based setting. A total of 395 names were acquired.

Postcards were mailed to the 395 participants asking them to complete a survey. Of the 395 SLPs who were contacted 35 responded, representing a 9% response rate. The postcard

provided instructions on how to access the electronic questionnaire. The participants were initially directed to a consent letter at the beginning of the questionnaire. At the end of the electronic consent letter, the participants were prompted to provide consent to participate; if consent to participate was selected the questionnaire was opened on Survey Monkey. The participants were not able to access the questionnaire without giving electronic consent.

Measures: Questionnaire

The questionnaire was created for the purpose of this study. The questionnaire was administered to SLPs in Pennsylvania employed in a school-based setting. The purpose was to collect information pertaining to their experiences working with bilingual clients, as well as their educational backgrounds in the areas of bilingual assessment and treatment. The questionnaire included questions on the background information of the participants to allow further analysis of the data if warranted. The questionnaire utilized "check all that apply" boxes, boxes for fill-in-the blank statements, multiple choice questions with only one viable answer, as well as open-ended questions. Open-ended questions were linked to close-ended questions on the survey and used as a supplement.

RESULTS

Both quantitative and qualitative data were collected. Primarily quantitative analyses were completed to answer the research questions with qualitative responses used to support the quantitative findings. The quantitative data were analyzed using descriptive and nonparametric procedures. The statistical procedure selected was matched to the research question with consideration given to the exploratory nature of this study.

Out of the 395 individuals contacted, 35 responded representing a return rate of 9%. The respondents worked in a variety of settings: early intervention (6%), preschool (3%), elementary school (84%), and high school (22%) settings. Several respondents worked at multiple settings and indicated accordingly on the survey. Twenty-four respondents (69%) had over 12 years of experience, whereas, 11(31%) had less than 12 years of experience in the field.

The participants who responded to the survey completed their graduate level training at the following schools in Pennsylvania: Bloomsburg University (n=6; 17%), Clarion University (n=9; 26%), Duquesne University (n=1; 3%), Edinboro University (n=3; 8.%), Indiana University of Pennsylvania (n=1; 3%), Penn State University (n=1; 3%), Temple University (n=2; 6%), University of Pittsburgh (n=3; 8%), and West Chester University (n-2; 6%). Seven respondents (20%) completed their graduate training outside of Pennsylvania. Respondents reported knowledge and use of languages other than English: 5 (14%) reported knowing no other languages, 25 (71%) reported knowing a few words or phrases in another language, 3 (8%) reported speaking 1 other language moderately well, 1 (3%) indicated fluent usage of 2 languages, and 1 (3%) indicated fluent usage in 3 or more languages. Refer to Table 1.

Number	35	
Work Setting	Early intervention	(6%)
	Preschool	(3%)
	Elementary school	(79%)
	High school	(18%)
Years of Work Experience	Over 12 years	(68%)
	Less than 12 years	(32%)
Graduate Level Training	Bloomsburg University	(17%)
	Clarion University	(26%)
	Duquesne University	(3%)
	Edinboro University	(8%)
	Indiana University	(3%)
	Penn State University	(3%)
	Temple University	(6%)
	University of Pittsburgh	(8%)
	West Chester University	(6%)
	Other	(20%)
Current Number of Bilingual Clients on Caseload		
	None	(54%)
	1-5	(34%)
	6-10	(3%)
	10+	(6%)
	Skipped	(3%)
Total Number of Bilingual Clients on Caseload During Career		
	None	(29%)
	1-5	(34%)
	6-10	(17%)
	11+	(20%)

Table 1. Participant information.

Nineteen respondents (54%) indicated no bilingual clients on their current caseload composition; whereas, 12 participants (34%) reported 1 to 5 bilingual clients, 1 participant (3%) reported between 6 to 10, and 2 participants (6%) reported 10 or more bilingual clients. During their career to date, 10 participants (29%) reported serving no bilingual clients, 12 (34%) reported providing services for 1 to 5 bilingual clients, 6 (17%) had worked with 6 to10 bilingual clients, and 6 (17%) had worked with 11 or more bilingual clients. One (3%) SLP responded that he/she had worked with 125 bilingual clients.

Twenty-two (63%) respondents had assessed a bilingual client during their career, whereas, 13 (37%) had not conducted a bilingual assessment. Of those who had conducted an assessment on a bilingual client, 16 (73%) had completed the assessment in English, and 5 (23%) had completed the

assessment in both English and the client's native language. No one reported conducting the assessment in only the first language of the client. When questioned about preparation to conduct the assessment, (n=25) 9 people (36%) responded that they felt prepared, 6 (24%) reporting being 'somewhat prepared', 5 (20%) reported they were 'not really' prepared, and 5 (20%) stated they were 'not prepared.' Of the respondents, 7 (20%) reported receiving graduate level coursework on bilingual assessment and treatment, and 28 (80%) did not recall coursework on the topic of bilingual therapeutic services. Of those who had received coursework, most reported it was not freestanding course but rather infused in another course, whereas, one person reported a freestanding course on the topic. Out of 35 responses, 4 (11%) had completed a continuing education course related to bilingual assessment and treatment, and 31 (89%) had not completed continuing education on the topic.

Chi-square tests of independence were performed to examine the relation between perceptions of preparation for assessment and treatment of bilingual services and graduate coursework and perceptions of preparation for assessment and treatment of bilingual services and continuing education. The relations were not significant. See Table 2.

The questionnaire asked SLPs about specific assessment measures, particularly standardized, norm-referenced instruments. Three respondents (9%) indicated standardized instruments were an accurate measurement of bilingual clients' skills, 15 (47%) indicated they were 'sometimes' accurate, and 14 (44%) reported they were 'not' an accurate measurement tool. The questionnaire also asked the SLP's, through a free response question, to provide positive factors regarding standardized evaluation tools for bilingual clients. There were several common themes that emerged in responses from the 16 respondents. Three (19%) suggested that standardized tests were reliable and equal. Some indicated standardized tests provide a baseline performance level for the second language (n=2; 13%), and allow the SLP to compare the client to English speaking students (n=2; 13%). SLPs noted some negative factors regarding standardized assessment for bilingual clients; including the tests are not a fair assessment and representation of skills (n=5; 28%), and lack cultural sensitivity (n=2; 11.%)

	df	$\chi 2$	*p*
Graduate Coursework	3	2.378	.498
Continuing Education	3	3.919	.270

Table 2. Perceptions of preparation for assessment and treatment of bilingual services and graduate coursework and continuing education. N = 25.

DISCUSSION

Out of the 35 respondents to the survey over half reported conducting bilingual evaluations and treatment. The SLPs who responded had a range of years of experience in the field across several different settings providing a small sample with diversity. A few factors of interest emerged from the results including the language in which SLPs assess bilingual clients, SLP's views on standardized measures when working with bilingual clients, and how this may relate to their graduate level training. Of those SLPs who had conducted a bilingual assessment 79% of them had conducted the assessment in English *only*. This is important to note because it is often difficult to determine the client's true language abilities or language deficits when assessed in a non-native language due largely to inherent variances in languages (ASHA, 2013). ASHA (2013) states that, "It is the ethical duty of all SLPs and audiologists to provide services with careful and respectful consideration and incorporation of the cultural and linguistic variables that have an impact on service delivery and efficacy." Differences in language are integral to the provision of culturally sensitive services, and individuals from other cultures may be at a disadvantage when assessed in *only* English; further, when SLP's assess only in English they have not provided a fair assessment or properly met the needs of the client with consideration given to preferred practice recommendations (ASHA, 2013). The scope of practice for the field of speech-language pathology clearly states that SLP's have a responsibility for "the provision of culturally and linguistically appropriate services…one aspect of providing culturally and linguistically appropriate services is to determine whether communication difficulties experienced by English language learners are the result of a communication disorder in the native language or a consequence of learning a new language" (ASHA, 2007).

With an ever-growing bilingual population, as well as the growing need for SLPs to assess these clients, do SLPs know how to adequately assess this population? It is considered best practice for SLPs to use alternative, nonbiased methods of assessment with their bilingual clients, such as language sampling and dynamic assessment (ASHA, 2013). However, not all SLPs have the necessary knowledge or skills to conduct a bilingual assessment with adequate cultural and linguistic sensitivity (ASHA, 2010; ASHA 2011) as also supported by the current results. Several SLPs indicated that a positive factor for the use of standardized, norm-referenced tests was gaining a baseline for the client's capabilities in English. However, as noted above, when you evaluate a client in *only* English you do not get a measure of their abilities in their native language. Due to a lack of understanding about the assessment of bilingual clients, those assessed in English *only* are likely to be misidentified. SLPs also stated that standardized measures are useful because they can then compare the client to English speaking students. Comparing bilingual persons to English speaking persons will not assist the SLP in evaluating whether or not they have a language disorder or provide information for

intervention planning. Since many SLPs responding to this questionnaire seemed unclear of preferred practice procedures identified by ASHA in the assessment of bilingual clients this likely implicates deficits in personal knowledge.

The questionnaire completed for the purpose of this project also contained an item regarding continuing education hours: *Have you ever completed continuing education on assessment and/or treatment of bilingual clients*? Thirty-five persons responded to the question with over 86% (n=31) reporting never taking a continuing education course or seminar on bilingual treatment or assessment. This implies that SLPs may be relying heavily on what they learn in graduate school to guide the assessment and treatment of bilingual clients. There appears to be a discrepancy between what SLPs are expected to do and what they are prepared to do. This discrepancy may be due to a lack of graduate level coursework provided to graduate related to bilingualism and linguistic diversity. Graduate programs may need to reevaluate their curriculum to ensure assessment and intervention is adequately covered with sufficient breadth and depth for linguistically and culturally diverse populations.

Although many SLPs reported involvement in bilingual assessment and treatment, only one of the respondents reported having coursework dedicated toward multicultural assessment and treatment during graduate training. The respondents had a wide range of experience working in the field as shown in Table 1. A decade or more ago, courses related to the provision of culturally and linguistically diverse services may not have been offered; however today, bilingualism and second language acquisition is more prevalent. Yet, graduate training programs may not offer dedicated coursework targeting assessment and intervention of culturally and linguistically diverse populations. From a brief review of accessible course directories, a few graduate programs in Pennsylvania offer freestanding courses such as West Chester University and Marywood University.

Most of the respondents to the survey reported having bilingual assessment and intervention immersed within another course during their graduate coursework; a limited number (n=2) could remember the specific course. Including the topic of cultural and linguistic diversity into another course may not be an effective means of teaching the subject with depth; McKenzie (2009) explained, "…infusion…limits access to that knowledge" (pg. 389). Graduate students are likely receiving a basic overview of the concepts related to assessment and treatment for bilingual individuals with limited instructional time dedicated to this topic Lack of coursework on bilingual assessment and treatment and the resulting lack of instructional breadth and depth the topic is given, is likely a reason why SLPs do not feel prepared to work with this ever-growing population.

Overall, the majority of SLPs had limited graduate level coursework or continued education training dedicated to bilingual assessment and treatment. Most graduate programs dedicate entire courses to disorder topics, such as stuttering; according to the U.S. Department of Health and Human Services, only about 1% of the population in the United States stutters (National Institute on Deafness and Other Communication Disorders, 2013), whereas at least 20% of the US population speaks a language other than English in the home (Ohlemacher, 2007). It could be hypothesized there is an equal chance if not higher that an SLP will work with someone with English as a second language during their career. Graduate programs may need to reevaluate their curriculum to ensure assessment and intervention is adequately covered with sufficient breadth and depth for linguistically and culturally diverse populations, especially since SLPs tend to rely on knowledge from graduate school. For example, future SLPs may need direct instruction on preferred assessment and treatment procedures including why standardized measures are not the favored assessment tool for use with bilingual populations.

Measurement error may limit this study. It is plausible that participants did not understand items on the questionnaire, did not provide the true answer, or other item wording effects interfered (Bethlehem, 2009). These measurement errors would result in differences between the obtained responses in the questionnaire and the true responses. The response rate for this survey was poor at 9% which resulted in a small sample size (n=35); several factors may be related to the low return rate.

First, the mailing occurred after the winter holiday season. Second, only 20 persons provided e-mail addresses, and due partially to funding constraints there were limited methods to follow-up with potential participants in an effort to facilitate questionnaire completion. Several persons reported technical difficulties with the online survey and were unable to complete the questionnaire. In the future, working collaboratively with state organizations to disseminate a similar questionnaire would be useful, thereby, circumventing the need to purchase potential participants. At the time of this pilot study, there was not a mechanism in place in Pennsylvania to utilize the state organization to disseminate this survey.

Future research should aim to increase the sample size and recreate this study with the assistance of state and national organizations to produce a broader data collection initiative. For example, a national initiative to document how academic institutions teach the principles and methods of assessment and treatment of culturally and linguistically diverse populations as aligned with preferred practices and ASHA policy documents could yield information pertinent to graduate training curriculums. In addition, it would be useful to investigate how graduate students respond to varying instructional methods including the impact of relevant practicum experiences.

CORRESPONDENCE
erineglundblom@gmail.com

REFERENCES

American Speech-Language-Hearing Association. (2003). *American English dialects* [Technical Report]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2004). Knowledge and skills needed by speech-language pathologists and audiologists to provide culturally and linguistically appropriate services [Knowledge and Skills]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2004b). Preferred practice patterns for the profession of speech-language pathology [Preferred Practice Patterns]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2007). *Scope of Practice in Speech-Language Pathology* [Scope of Practice]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2008). 2008 Schools Survey summary report. Available from http://www.asha.org/research/memberdata/schoolssurvey/.

American Speech-Language-Hearing Association. (2010). 2010 Schools summary report. Available from http://www.asha.org/research/memberdata/schoolssurvey/.

American Speech-Language-Hearing Association. (2010c). Roles and responsibilities of speech-language pathologists in schools [Professional Issues Statement]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association. (2011). *2011 Membership survey. CCC-SLP survey summary report.* Available www.asha.org/research/memberdata/2011-membershipsurvey.

American Speech-Language-Hearing Association. (2011). Cultural competence in professional service delivery [Professional Issues Statement]. Available from www.asha.org/policy.

American Speech-Language-Hearing Association (2012). Dynamic assessment: Basic framework. Retrieved http://www.asha.org/practice/multicultural/issues/framework.htm.

Annett, M. (2002). New Mexico clinicians seek state recognition for bilingual skills. *ASHA Leader.* Retrieved http://www.asha.org/Publications/leader/2002/020402/020402b.htm

Bethlehem, J. (2009). *Applied survey methods: A statistical perspective.* Hoboken, New Jersey: Wiley.

Bureau of Labor Statistics, U.S. Department of Labor, *Occupational Outlook Handbook, 2008-09 Edition,* Speech-Language Pathologists, on the Internet at http;//www.bls.gov/ooh/healthcare/speech-language-pathologists.htm.

Caesar, L., & Kohler, P., (2007) The state of school-based bilingual assessment: Actual practice versus recommended guidelines. *Language, Speech, and Hearing Services in Schools.* 38, 190-200.

Dolloghan, C., & Horner, E. (2011). Bilingual language assessment: A meta-analysis of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research Vol. 54.* 1077-1088

Geisler, E. (2010). Difference or disorder? Researchers develop tool to identify bilingual children with true language disorders. Retrieved http://www.utexas.edu/features/2010/09/27/language-2/

Grosjean, F. (2012). Bilinguals in the United States of America. *Psychology Today.* Retrieved http://www.psychologytoday.com/node/95889.

Gutierrez-Clellen, V., Adelaida Restrepo, M., Bedore, L., Pena, E., & Anderson, (2000). R. Language sample analysis in spanish-speaking children: Methodological considerations. *Language, Speech, and Hearing Services in Schools. 32,* 88-98.

Kapantzoglou, M., Retrepo,M. A., & Thompson, M. S. (2012) Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools. 43,* 1-96,

Kritikos, E., (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. *American Journal of Speech-Language Pathology, 12.* 73-91.

Limbos, M. M., & Geva, E. (2001). Accuracy of teacher assessments of second-language students at risk for reading disability. *Journal of Learning Disabilities, 34*(2), 136-151.

Ohlemacher, S. (2012) 20 percent of people living in U.S. speak language other than English at home. *The Post and Courier.* Retrieved March 17, 2013 from http://www.postandcourier.com/article/20070912/PC1602/309129945

Peña, E. D., Bedore, L. M., & Rappazzo, C. (2003).) Comparison of Spanish, English, and bilingual children's performance across semantic tasks. *Language, Speech, and Hearing Services in Schools*, *16*(13), 5-16.

Peña, E. D. & Bedore, L. M. (2011). It takes two: Improving assessment accuracy in bilingual children. *The ASHA Leader, 16(13),* 20-22.

Pennsylvania Department of Education. (2013). English as a second language. Retrieved March 3, 2013 from http://www.education.state.pa.us/portal/server.pt/community/english_as_a_second_language/7529.

Tzuriel, D. (2001). *Dynamic assessment of young children.* New York, NY: Kluwer Academic/Plenum

U.S. Department of Health and Human Services. National institutes on deafness and other communication disorders (NIDCD). Retrieved March 21, 2013 from https://www.nidcd.nih.gov/health/voice/pages/stutter.aspx.

AUTHOR INFORMATION

Koren Beardshall, B.S., is a graduate of the undergraduate program in Communication Sciences and Disorders at Clarion University of Pennsylvania. During her time at Clarion, Koren was a student in the Honors Program, a member of the Communication Science and Disorders Research Collaborative, as well as NSSHLA member. Koren is also a member of the E. G. honors societies Phi Eta Sigma and Phi Kappa Phi. Koren is continuing her graduate education at Duquesne University.

Erin Lundblom, Ph.D., CCC-SLP, received her doctoral degree from Florida State University. She has been an Assistant Professor at Clarion University and is now a faculty member at The University of Pittsburgh. Her personal research interests include the application of research to practice including evidence-based practice in student supervision and school-based language literacy.

INFINITIVAL CLAUSES IN CHILDREN WITH TYPICAL AND LATE LANGUAGE EMERGENCE: SUPPORTING A DIMENSIONAL ACCOUNT OF LANGUAGE DELAY

Celeste Domsch
Texas State University
Corrin Richels
Old Dominion University
Maria Resendiz
Texas State University

ABSTRACT

Children who do not speak single words by 2 years of age have been labeled as having late language emergence (LLE). While the majority of children with LLE recover by school-age, it has been argued that they often still perform below the level of their typical peers for specific linguistic skills. In this case, speech-language pathologists (SLPs) should consider language skills as varying along a dimension, rather than as simply impaired vs. unimpaired. To examine the dimensionality of language skill, this study compared infinitival clause production in 22 school-age children with and without LLE. The infinitive clauses were: catenatives, such as *gonna*; *let us* + verb, typically produced as *let's*; unmarked infinitives such as *make it go*; and simple infinitives such as *We want to run*. The 22 participants included 11 with typical development and 11 with a history of LLE, sampled in a conversational context at 8-years of age. Analysis indicated that the groups did not statistically differ for use of the four types of infinitival clauses. However, the LLE group did use fewer simple infinitives, offering support for a dimensional model of language development.

KEY WORDS

Late Language Emergence
Infinitive Clause Production, Syntactic Development
Dimensional Account of Language Delay

INTRODUCTION

Children with late language emergence (LLE) are those for whom the onset of expressive language is delayed to 24 months or later (Zubric, Taylor, Rice, and Slegers, 2007). This delay in expressive language has been operationalized as a productive vocabulary of fewer than 50 words at 2 years, or a score in the lowest 10th percentile on a standardized parent questionnaire regarding vocabulary size. Children with LLE have normal hearing and nonverbal IQ scores, and no obvious neurological impairments. Children with known medical conditions or who are bilingual have typically been excluded from studies of LLE.

Definitions of LLE have shown some variability with regard to age and expressive vocabulary size, with production of two-word combinations as a potential third factor to consider. Some have restricted use of the LLE diagnosis to children at 24 months of age (Zubric et al, 2007), while others have included children up to 35 months of age (Girolametto, Wiigs, Smyth, Weitzman, and Pearce, 2001). For vocabulary size, the commonly suggested cutoff of fewer than 50 words has been implemented by several researchers. Paul and Smith (1993), for example, reported a mean vocabulary size of 27.7 words, with no range given. Similarly, Rescorla (2009) described a group of children with LLE who produced an average of 24.54 words, with a range = 5-131, in contrast to a comparison group of typically developing (TD) children who produced an average of 235.17 words, with a range = 27-319. Both of these studies employed a parent questionnaire to measure expressive vocabulary size. Given the ranges for vocabulary size just noted, both studies included children with vocabularies larger than 50 words in their LLE groups. For these children, both studies reported that a lack of two-word combinations was taken as diagnostic of LLE.

Other studies (e. g., Girolametto et al., 2001) have used a percentile score for vocabulary production as their diagnostic indicator for LLE. The *MacArthur Communicative Development Inventory* (CDI) (Fenson, Dale, Reznick, Thal, and Bates, 1993) is a 680-word parent checklist, which yields a percentile score for vocabulary production for children between the ages of 16-30 months. Thal, Tobias, and Morrison (1991) used a score in the lowest 10th percentile on an early version of the CDI as their cutpoint for LLE, while Girolametto et al. used the lowest 5th percentile.

Other, larger studies have employed yet a different criterion, beyond expressive vocabulary size in words or a percentile score on a checklist. Zubrick et al. (2007), in an epidemiologically-ascertained sample of 1,766 children, defined LLE as a score of -1.0 SD or below on a communication subscale that asks whether a child points to pictures and body parts, follows simple directions, names objects, combines words, and/or uses early-developing personal pronouns. Using that varied tasks and this cutoff, 13.4% of the sample was diagnosed with LLE.

Children with LLE at Kindergarten Entry: Short-Term Outcomes

Many children with LLE exhibit significant growth in vocabulary during the time between diagnosis and school entry, with a majority of them scoring in the average range for language skills at kindergarten (Paul, 1996). Rescorla, Roberts, and Dahlsgaard (1997), for example, reported that children diagnosed with LLE between 24 and 31 months scored in the average range for single-word vocabulary by 3-years, indicating significant growth in what was for some only a five-month period. Likewise, Paul (1993) reported that 37 children with LLE did not differ from typically-developing children for scores on a standardized vocabulary test administered at age 3.

Despite this potential short-term growth in vocabulary, some children with LLE remain below average for syntactic achievement. Many of the bound morphemes of English are acquired between ages 2 and 4, which results in increases to children's mean-length-of-utterance (MLU) during this time. Rescorla, Roberts, and Dahlsgaard (1997) analyzed conversational samples for MLU in 34 children who were diagnosed with LLE between 24-31 months, and then seen for a follow-up visit at 36-months. The children with a history of LLE had an average MLU *z* score of -1.51, indicating that they were either failing to use bound morphemes at a rate similar to peers, or that they were failing to combine words at similar rates, or both. The same study used the *Index of Productive Syntax* (IPSyn) (Scarborough, 2010) which evaluates noun and verb phrase elaboration, the use of questions and negation, and overall sentence structure. The children with a history of LLE performed even lower than they did for MLU, with an average IPSyn *z* score of -2.21. It appeared that the early language delay that may have improved or even resolved relative to single-word acquisition had not improved for more sophisticated language tasks.

In regard to syntactic delays, Paul [9] found that 60% of her sample of children with LLE scored below the 10th percentile on the *Developmental Sentence Score* (DSS) (Lee, 1974), which assesses use of indefinite and personal pronouns, main and secondary verbs, negation, and the use of questions. Ellis Weismer (2007) compared language outcomes at age 5.5-years for 40 children with LLE with those of 43 TD peers. Although the children with LLE scored in the average range on a standardized language test, their mean scores were significantly below those of their TD peers, particularly in sentence imitation, a task which requires the use specific bound morphemes and often complex syntactic structures.

Children with LLE in the Early Grades: Longer-Term Outcomes

The longer-term outcomes for this population will affect their educational placements and their academic progress. As Paul (1996) reported, the majority moved into the average range at kindergarten, while some continued to require clinical attention. It is important to note that even children scoring in the average range for language skills in kindergarten may fall

out of the average range once vocabulary demands increase, once syntactic structures become more complex, and once independent reading is expected, such as in second and third grades (Nippold, 2007).

To address the question of longer-term outcomes, Rice, Taylor, & Zubrick (2008) examined syntax in conversation in 7-year-olds with and without a history of LLE. Results indicated that those with a history of LLE demonstrated significantly lower MLU in morphemes, as well as increased errors on a number of verb structures (i.e., marking of past tense, use of copula and auxiliary "be," and use of auxiliary "do" etc.). In an even longer-term study, Rescorla (2009) found that teens with a history of LLE tended to score lower on standardized tests of grammar than teens without such history, and argued that slow early language development may reflect a predisposition to lower linguistic performance over time. In a recent review paper, Rescorla reported that both small and large *n* epidemiological studies examining long-term outcomes in children with LLE have supported the dimensional account of language delay, "whereby late talkers and typically developing peers differ quantitatively on a hypothetical language ability spectrum" (Rescorla, 2013, p 141).

Infinitival Clause Development in TD Children and Those with SLI

In a concise review of the development of infinitives in TD children, Eisenberg (2004) notes that infinitive use begins around 2 years of age with a restricted set of verbs (e.g., *go*, *want*, *have*), and that they emerge as catenatives (e.g., I *wanna* eat, He's *gonna* jump, We *hafta* wash our hands) or as being unmarked (e.g., *Help me do it*). Some researchers have labeled *let* clauses (e.g., *Let's eat*) as unmarked infinitives (Steffani, 2007), while others have separated *let* into an infinitive category of its own (Scheule, 2009), likely due to the frequency of its use in young TD children (e.g., Let's *pretend*). Unambiguous use of the infinitive marker *to* (e.g., *I need to move that*) occurs once MLU has reached approximately 3.5 (Bloom, Tackeff, and Lahey, 1984) and is labeled as a simple infinitive. The simple infinitive category typically excludes verbs that are usually produced as catenatives, including *gonna*, *wanna*, and *hafta* as noted above.

Regarding infinitive clause development in children with specific language impairment (SLI), findings have been equivocal. Marinellie (2004) examined use of infinitives in conversation for 10-year-old children with and without SLI, and reported no significant differences. In contrast, Leonard, Eyer, Bedore, and Grela (1997) reported lower use for the infinitive *to* marker on a sentence completion task as compared to both age- and MLU-matched peers. Likewise, Eisenberg (2003) reported that children with SLI produced infinitival object complements with fewer verbs than did TD children of the same age or even younger. However, when Eisenberg (2004) compared 5-year-olds with SLI to typical children ranging in age from 3- to 5-years using an elicited production task for infinitival complements, she found that all eight children with SLI demonstrated production of infinitives with a variety of main verbs, and that only one child of the eight with SLI omitted the *to* marker more than once. Given that performance in an elicited task was relatively strong, Eisenberg argued that "the limited production of infinitives in conversation may in part reflect a problem with mobilizing syntactic knowledge (a performance issue) rather than a lack of knowledge per se (a competence issue)" (Eisenberg, 2004, p 319). Thus, it is possible that children with SLI have an "infinitive structure-finding" problem, in the same way that some children demonstrate receptive knowledge of specific vocabulary items yet cannot readily produce those words in conversation and so are labeled as having word-finding difficulties. Arndt and Schuele (2012) also studied infinitival complement use in children with SLI as compared to younger, MLU-matched children with typical language development. Comparison of infinitives used in spontaneous language samples found no difference in the number of infinitival complements or the number of different complement-taking verbs, but the children with SLI were significantly less likely to include the infinitival marker *to*, which Arndt and Schuele took as evidence that these children were experiencing "difficulty with the specific grammatical requirement of infinitival clauses" (Arndt and Schuele, 2012, p. 1).

The current study provides a test case regarding long-term outcomes in LLE by examining the specific use of four infinitival clause types (e.g., catenatives, *let's*, unmarked infinitives, simple infinitives) in a conversational context, where production is not intentionally modeled. Infinitive use in TD children and those with SLI has been examined in both conversational language and elicited tasks, as noted above, but has not, to the best of our knowledge, been specifically tabulated in school-age children with LLE. Whether school-age children with a history of LLE perform more like their TD peers or more like children with SLI for these tasks will shed light on the extent to which language development in those with a history of early delay remains weaker than expected, thereby supporting the dimensional model of Rescorla (2013). The specific questions are as follows:

Do 8-year-olds with and without a history of LLE use catenatives at similar rates in conversational samples?

Do 8-year-olds with and without a history of LLE use *let's* at similar rates in conversational samples?

Do 8-year-olds with and without a history of LLE use unmarked infinitives at similar rates in conversational samples?

Do 8-year-olds with and without a history of LLE use simple infinitives at similar rates in conversational samples?

METHODS

This study was approved by the Institutional Review Board of Texas State University-San Marcos. Informed consent was obtained from the parents of all participants, and assent was obtained from all children in the study.

Participants

Twenty-two children participated in this study, 11 with a history of LLE and 11 with a history of TD. The children in the LLE group were participants in a previous study of early vocabulary growth (Roid & Miller, 2001). That study included 20 children with LLE, who were recruited through newspaper advertising, fliers distributed at daycare centers, and word-of-mouth. The first author attempted to locate all 20 children approximately five years after completing the initial study, but families could not be located (n = 3), or had moved out of state (n = 2), or they declined participation (n =2), or their data were lost (n =2). As a result, this study included a group of 11 children who had LLE as toddlers. At intake, these 11 children had a mean age of 29.7 months (SD = 4.4, range = 24-39).

LLE was diagnosed using the following inclusion criteria. First, children exhibited a reduced vocabulary size, defined as a score below the 10th percentile for their ages on the CDI, as in Thal and colleagues (1991). For the children older than 30 months, vocabulary production scores were below the 10th percentile for 30 months. The mean vocabulary size for these 11 children was 63.4 words (SD = 56, range = 8-188). Second, they exhibited an average nonverbal IQ score on the Brief IQ subtest of the *Leiter International Performance Scale-Revised* [26] with a mean score of 104.00 (SD = 10.51, range = 85-117). Third, English was the only language spoken in the home. Finally, parents reported no history of hearing impairment, autism, or any other neurological disorder. The average level of maternal education was 14.45 years (SD = 2.77). One child was female, and the rest were male. Of the 11 children, 9 were Caucasian, one was African American, and one was Asian-American. At the time of this follow-up study, the 11 children who had LLE as toddlers now had a mean age of 8.6 years (SD = 0.36). All were in mainstream classrooms in public schools in central Texas at the time of this study.

The rate of attrition (11 children with LLE found out of an original n of 20) raises the possibility that the children "lost" to follow-up might have differed from the children "found" for follow-up. Thus, the 9 children lost for follow-up were compared to the 11 children found for follow-up for gender and maternal education in years, as well as their toddler measures of nonverbal IQ and vocabulary size. A 2 x 2 contingency table analysis showed that the proportion of the "lost" group who were males (67%, n = 6) was significantly lower than the proportion of the "found" group who were males (91%, n = 10), χ^2 (1, N = 20) = 1.82, p < .05. Three females were lost to follow-up from the original group, and only one female was retained. Comparisons for the other variables were calculated using independent samples t – tests with equal variances assumed. No significant differences were noted between the two groups for maternal education (t(18) = .315, p = .756), nonverbal IQ (t(18) = .212, p = .834), or vocabulary size t(18) = -.544, p = .593).

The 11 children in the TD group were recruited through contacts at Texas State University (n = 4 children) and an elementary school in the Austin Independent School District (n = 7). Their mean age was 8.5 years (SD = 0.20), and their average level of maternal education was 15.73 years (SD = 2.20). All 11 children were Caucasian. Six were male and five were female. All children were reported by their parents to be functioning on grade level and receiving no special education services. They learned to talk at the expected age, by parent report, and had never received speech-language therapy.

Data Collection and Analysis

Participants were visited in their homes by the first author, and engaged in ten minutes of conversation, which was videotaped. Topics included school, family members, holidays, and favorite activities. It should be noted that conversational language samples are, by their nature, uncontrolled for content and/or syntactic difficulty. Video samples were transcribed by graduate students. Transcripts were segmented into C-units, defined as an independent clause plus any modifiers. C-units could also include coordinated clauses, defined as one main clause plus one additional clause that was introduced with *and*, *but*, or *or*, following procedures in Marinellie (2004).

Once transcription and segmentation into C-units was complete, the samples were coded for the four infinitive structures of interest (Schuele, 2009). C-units containing *gonna*, *wanna*, *gotta*, *sposta*, and *hafta* were coded as catenatives, and any use of *let's* was coded for that category. Production of unmarked infinitives was coded for C-units containing *make*, *help*, and *watch* that did not include the infinitival marker *to*. Simple infinitives were coded for C-units that included the infinitival marker *to*, followed by a verb. Reliability for identification of the four infinitival clauses was 91% for six randomly-selected transcripts (three LLE and three TD) across all four clause types when coded separately by the graduate student transcribers as compared to the first author. Disagreements in coding were resolved through discussion.

RESULTS

Gender

A 2 x 2 contingency table analysis showed that the proportion of the LLE group who were males (91%, n = 10) was not significantly greater than the proportion of the TD group who were males (55%, n = 7), Pearson χ^2 = (1, N = 22) = 3.67, p = .056. Because this result was not statistically significant, the gender of the participants was not included as an independent variable.

Levene's Test for Equality of Variances

Levene's test was calculated to ensure that the dependent variables did not violate the assumptions of homogeneity of variance necessary for MANOVA. A statistically significant result indicates that the equality-of-variance assumption is violated (Green and Salkind, 2011). Results for the dependent measures indicated that this assumption for the dependent measures (e.g., use of catenatives, *let's*, unmarked infinitives, and simple infinitives) was not violated for homogeneity of variance. Therefore all variables were included in the MANOVA.

Demographic Variables

Results of an independent samples *t*-test with age (unequal variances assumed) and maternal education (equal variances assumed) as the dependent variables and group membership (LLE vs. TD) as the independent variable found no significant differences between the groups for age, (LLE M = 8.66, SD = .38, TD M = 8.58, SD = .20; $t(15.2)$ = .588, p = .565) or maternal education (LLE M = 14.45, SD = 2.77, TD M = 15.73, SD = 2.20; $t(20)$ = -1.19, p = .25). Due to this non-significant finding, no further corrections were made for age or maternal education in the analyses.

Infinitival Clause Use

A MANOVA was computed to determine the effect of language history (e.g., LLE vs. TD) on the use of catenatives, *let's*, unmarked infinitives, and simple infinitives. Dependent variables were the number of catenatives, *let's*, unmarked infinitives, and simple infinitives produced by each participant. The independent variables were talker group membership as a child diagnosed with LLE or as a child with TD. Results of the MANOVA indicated no main effect for the dependent variables and talker group, Wilks' Λ = .827, $F(1,20)$ = .891, p = .490. Table 1 contains the means and the standard deviations on the dependent variables for the two talker groups.

	LLE Control
Catenatives	1.10 (*1.58*) 0.727 (*1.79*)
Let's	0.364 (*0.674*) 0.818 (*2.40*)
Unmarked infinitives	0.910 (*1.14*) 1.09 (*0.944*)
Simple infinitives	10.5 (5.68) 15.09 (6.38)

Table 1. Means (*standard deviations*) for dependent variables by talker group (LLE n = 11; and TD n = 11.)

Follow-up ANOVAs were calculated with significance level corrected using the Bonferroni correction to $p < .012$ (4 comparisons). Results indicated no significant differences between the groups for the use of catenatives, ($F(1, 20)$ = .255, p = .619, partial η^2 = .013), *let's* clauses, ($F(1, 20)$ = .365, p = .552, partial η^2 = .018), unmarked infinitives, ($F(1, 20)$ = .167, p = .552, partial η^2 = .008), and simple infinitives ($F(1, 20)$ = 3.24, p = .087, partial η^2 = .139).

The typically-developing group and the LLE group did not differ for use of catenatives, unmarked infinitives, or the use of *let's*. All three were relatively low-frequency structures, occurring on average fewer than three times per sample. Although the difference between groups was not significant, the typically-developing group used more simple infinitives in conversation than did the LLE group (M = 15.1 productions vs. M = 10.5 productions, respectively). It is worth noting that simple infinitives emerge once MLU has reached 3.5, so reduced usage at age eight by the LLE group hints at a differing level of proficiency with this complex structure.

DISCUSSION

First, the finding that catenatives, *let's,* and unmarked infinitives are relatively infrequent in conversation at age eight is consistent with results from Marinellie [20], who reported a range from 0-8 uses for catenatives and *let's* clauses combined for TD 10-year-olds in conversational samples (M= 3.00, SD = 2.2), as compared to a range of 0-7 uses in 10-year-olds with SLI (M = 2.73, SD = 2.73). Catenatives are restricted in number by definition, as not every verb can form a catenative. *Let's* may occur more frequently in the conversation of younger children (e.g., *Let's pretend*) than it does in that of older children. Unmarked infinitives, like catenatives, can only occur with specific verbs (e.g., *make, help*), which may also be unlikely to occur in a conversational setting between a school-age child and an examiner, where there are no toys or manipulatives in use.

Second, while the use of simple infinitives did not statistically differ between the TD and LLE groups, the M number of productions was 15.09 vs. 10.5, respectively. These results are consistent with those reported by Marinellie (2004), who combined simple and unmarked infinitives into a single category, and found no significant differences in use between a typical group and a group with SLI. In that study, the TD group used a M of 9.13 (SD = 4.54) infinitive clauses, while the group with SLI used a M of 6.73 (SD = 4.09) infinitive clauses. These results are also consistent with reports that children with LLE score in the average range for standardized tests of language skill (Ellis Weismer, 2007; Resorla, 2009), though their specific scores are often lower than those of their TD peers.

It is noteworthy that none of the long-term studies of children with LLE have reported their participants earning significantly higher scores than TD peers. Instead, outcomes indicate either no significant differences (though scores may be lower for the LLE group), or the LLE group does indeed score significantly lower. The results of this study lend further support for a dimensional account of language delay (Resorla, 2013) and suggest that SLPs treating school-age children would be well-advised to probe the language development history even of children who are well past first words, and to adopt a more nuanced model of assessment than a binary impaired vs. unimpaired model.

Finally, it should be noted that one of the advantages of conversational language sampling (i.e., its ecological validity, as the child is allowed to choose the topic and direct the interaction), which has made it a standard language assessment task for years, is also a disadvantage in that children make very

different choices about what to discuss, which then influences the syntactic nature of their language. In reviewing the transcripts of the children in this study, it was apparent that some children chose to discuss family members, teachers, and favorite foods, and that these topics may have provided fewer opportunities for the emergence of complex syntax. Other children chose to describe a favorite vacation, which led to use of a narrative structure that included elements of story grammar (e.g., initiating events, conflict resolution, etc.), which provided more opportunities for complex sentences though use of conjunctions such as *because*. Still other children chose to explain a favorite game or sport, which has been described as an expository task and one which elicits perhaps the most use of complex syntax (Nippold, 2007).

Thus, it appears that children respond differently to the same conversational prompts, though it was outside the scope of this study to perform an ethnographic analysis of topic selection and any potential influence of it on syntactic complexity.

Limitations

The attrition rate from the Domsch and Camarata original study (2008) is one potential limitation of this study, though it should be noted that the children with LLE who were found for follow-up did not differ from the children with LLE who were lost to follow-up for vocabulary size, nonverbal IQ, or maternal education. The ethnicity of the participants in this study was relatively homogenous, with a large proportion being Caucasian. One of the often-used diagnostic criteria for LLE, which was employed in this study, has been exposure to English-only in the home, which obviously eliminates the participation of bilingual families of every ethnicity. Finally, the size of the LLE group is small (n = 11), which contributes to the possibility of a Type II error. It is possible that differences between groups would have emerged in a larger sample.

Future Research

One issue in the study of infinitival complements is that previous research studies have grouped clauses differently, with some including unmarked infinitives and simple infinitives in the same category while others separate them. These differing patterns of data analysis make it difficult to compare results across studies. In addition, most studies have employed a single task to measure infinitive production, which in some cases has been conversational speech, while in others infinitives have been elicited in single sentences. Future studies should include multiple methods for assessing infinitive production, as it would be useful to compare conversational vs. elicited production in a single large sample. Finally, the type of spontaneous language task administered should include expository discourse, in addition to conversation. Expository discourse is the main type of discourse that school-age children encounter at school, which they are required to both comprehend and produce. Expository discourse tasks have also been shown to yield more syntactically complex language than conversational samples (Nippold, 2007).

ACKNOWLEDGMENTS

The authors acknowledge Dr. Melanie C. Schuele for use of her Complex Syntax Coding Manual. They are grateful to Crista Lindsey, M.S., CCC-SLP for her coding of infinitival clauses. They are also grateful to the families who participated in this study.

CORRESPONDENCE

cd23@txstate.edu

REFERENCES

Arndt, K. B., & Schuele, C. M. (2012). Production of infinitival complements by children with specific language impairment. *Clinical Linguistics and Phonetics*, 26, 1-17.

Bloom. L., Tackeff, J., & Lahey, M. (1984). Learning to in complement constructions. *Journal of Child Language,* 11, 391-406.

Domsch, C. & Camarata, S. (2008) Children with late language emergence: Effects of maternal education and language use. *Tejas: Texas Journal of Audiology and Speech-Language Pathology*, 31, 7-14.

Eisenberg, S. (2003). Production of infinitival object complements in the conversational speech of 5-year-old children with language impairment. *First Language*, 23, 327-341.

Eisenberg, S, (2004). Production of infinitives by 5-year-old children with language impairment on an elicitation task. *First Language*, 24, 305-321.

Ellis Weismer, S. (2007). Typical talkers, late talkers, and children with specific language impairment: A language endowment spectrum? In R.Paul (Ed.), The influence of developmental perspectives on research and practice in communication disorders: A Festschrift for Robin S. Chapman, pp. 83-102. Mahwah, NJ: Erlbaum.

Fenson L., Dale, P., Reznick, S., Thal, D., & Bates, E. (1993). MacArthur Communicative Development Inventories. San Diego, CA: Singular.

Girolametto, L., Wiigs., M, Smyth, R., Weitzman, E., & Pearce, P. S. (2001). Children with a history of expressive vocabulary delay: Outcomes at 5 years of age. *American Journal of Speech-Language Pathology,* 10**,** 358-369.

Green, S. B. & Salkind, N. J. (2011). Using SPSS for Windows and Macintosh: Analyzing and understanding data (6th Ed.). Upper Saddle River, NJ: Pearson Education.

Lee, L, (1974). Developmental sentence analysis. Evanston, IL: Northwestern University Press.

Leonard, L., Eyer, J. A., Bedore, L. M. , & Grela, B. G. (1997). Three accounts of the grammatical morpheme difficulties of English-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 40, 741-753.

Marinellie, S. (2004). Complex syntax used by school-age children with specific language impairment in child-adult conversation. *Journal of Communication Disorders*, 37, 517-533.

Nippold, M. (2007), Later language development: school-age children, adolescents, and young adults. Austin, TX: Pro-Ed, Inc.

Paul, R., (1993). Patterns of development in late talkers: preschool years. *Journal of Childhood Communication Disorders*, 15, 7-14.

Paul, R. (1996). Clinical implications of the natural history of slow expressive language development. *American Journal of Speech-Language Pathology*, 5, 5-21.

Paul, R, & Smith R. L. (1993) Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech and Hearing Research,* 36, 592-598.

Rescorla, L. (2009). Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *Journal of Speech, Language, and Hearing Research*, 52**,** 16-30.

Rescorla, L. (2013). Late talkers: Do good predictors of outcome exist? *Developmental Disabilities Research Reviews*, 17, 141-150.

Rescorla, L., Roberts, J., & Dahlsgaard, K., (1997). Late talkers at 2: Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, 40, 556-566.

Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research*, 51, 394-407.

Roid, G. & Miller, L. (2001). Leiter International Performance Scale-Revised. Los Angeles, CA: Western Psychological Services.

Scarborough, H. (1990). Index of productive syntax. *Applied Psycholinguistics,* 11, 1-22.

Schuele, C, M. (2009). Complex syntax coding manual (Unpublished). Vanderbilt University: Nashville, TN.

Steffani, S. (2007). Identifying embedded and conjoined complex sentences: Making it simple. *Contemporary Issues in Communication Sciences and Disorders*, 34, 44-54.

Thal, D., Tobias, S., Morrison, D. (1991). Language and gesture in late talkers: A one-year follow up. *Journal of Speech and Hearing Research*, 34, 604-612.

Zubrick, S. R, Taylor, C. L., Rice, M. L., & Slegers, D.W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research*, 50, 1562-1592.

AUTHOR INFORMATION

Celeste Domsch (Ph.D., Vanderbilt University) is an Assistant Professor at Texas State University with interests in lexical development in late talkers and longitudinal research with this population. Corrin Richels (Ph.D. Vanderbilt University) is an Assistant Professor at Old Dominion University with an interest in child language development and stuttering disorders. Maria Resendiz (Ph.D. The University of Texas-Austin) is an Assistant Professor at Texas State University with an interest in the narratives performed by children with cultural and linguistic differences.

AGREEMENT AMONG COMPUTERIZED SPEECH LAB, PRAAT, AND WAVE SURFER IN DETERMINING FORMANT FREQUENCIES

Elizabeth U. Grillo
Ralph Magliano
Jodi DiSario
West Chester University

ABSTRACT

The purpose of this study was to determine whether or not agreement existed among the Computerized Speech Lab (CSL), Praat, and WaveSurfer in measuring the first three formant frequencies of the point vowels (/i/, /μ/, /æ/, /ɑ/) for ten participants with normal speech and voice. Two measurement techniques were used to identify the formant frequencies: 1) analysis of the three formants by hand at a specific time interval and 2) the average of each of the three formants reported by the software program. Results suggested that when measured by hand at a specific time interval, CSL and WaveSurfer did not agree in formant value analysis, whereas Praat did agree with the other two software programs. When both measurement techniques were compared across CSL and Praat, the formant value analysis did not agree.

KEY WORDS

Acoustic
Computerized Speech Lab
Praat, WaveSurfer
Formant Frequencies, Formants

INTRODUCTION

The Scope of Practice in Speech-Language Pathology describes a framework for clinical practice that includes "using instrumentation to observe, collect data, and measure parameters of communication and swallowing or other upper aerodigestive functions" (ASHA, 2007). Consistent with this framework, speech-language pathologists (SLP) use specialized acoustic analysis software programs to convert recorded speech into interactive visual displays for assessment and treatment of speech and voice. The various types of acoustic measurements allow for a more precise diagnosis, provide evidence for interventions, and present visual feedback for patients in therapy. The validity and reliability of such visual information is essential to accurate clinical assessment and treatment.

Three spectrographic software programs were assessed in the current study, Computerized Speech Lab (CSL), Praat, and WaveSurfer. CSL (model 4150) is available for purchase from KayPentax (Montvale, NJ). The Department of Communication Sciences and Disorders at West Chester University owns a license to run the CSL and uses it for research and clinical purposes. The latest versions of Praat and WaveSurfer are free and available to download at the following websites:

Praat
http://www.fon.hum.uva.nl/praat/
WaveSurfer
http://www.speech.kth.se/WaveSurfer/

CSL, Praat, and WaveSurfer were chosen for comparison because they are three widely used programs.

Formant frequencies are the resonant frequencies of the vocal tract and are constantly changing due to the shape of the vocal tract (Fant, 1960). The focus of the current study involved the first three formant frequencies of the point vowels (/i/, /μ/, /æ/, /ɑ/). The point vowels were chosen for analysis because they represent the extremes of tongue height and tongue advancement which ultimately influences the formant frequencies; therefore, of all the vowels in English, the point vowels demonstrate the clearest differences in the formants allowing for ease of measurement (Ladefoged, 1996; 2006; Ladefoged, Harshman, Goldstein, & Rice, 1978). Formant one is related to tongue height and formants two and three are related to tongue advancement (Ladefoged, 1996; 2006; Ladefoged, Harshman, Goldstein, & Rice, 1978). All of the software programs allow for measurement of the formants by hand at a specific time interval. Only two of the programs produce an average value of the entire formant, CSL and Praat. Because two measurement techniques were used for formant value analysis, both of the techniques were assessed for agreement among the software programs.

To date, no prior study has investigated whether or not agreement exists among CSL, Praat, and WaveSurfer for analysis of the first three formant frequencies during production of the point vowels. One prior study compared various acoustic measures by two different programs, Dr. Speech (DRS, Tiger Electronics, Neu-Anspach, Germany) and CSL (Smits, Cueppens, & De Bodt, 2005). A group of 120 participants with normal voices produced /ɑ/ at a comfortable pitch and loudness. The following parameters were selected for analysis: fundamental frequency (F_o), variation of F_o (F_oSD), jitter, shimmer, and noise to harmonics ratio (NHR). A significant correlation was found for F_o, NHR, and shimmer. The correlation for jitter and F_oSD was weak. DRS and CSL did agree for measurements of F_o, NHR, and shimmer, but did not agree for measurements of jitter and F_oSD. Based on the results in Smits and colleagues (2005), the literature does suggest a difference in agreement between analysis programs for acoustic measures; however, no study to date has analyzed formant frequency analysis of the point vowels across CSL, Praat, and WaveSurfer.

The purpose of this study, therefore, was to compare three voice analysis software programs, CSL, WaveSurfer, and Praat for formant frequency analysis of the four point vowels. Specifically, there were three experimental questions.

1) Will CSL, Praat, and WaveSurfer agree in formant frequency analysis of the four point vowels when measured by hand at 2.5 seconds?

2) Will the means of the formant values for the four point vowels agree when measured by hand at 2.5 seconds and when averaged by the programs across CSL and Praat?

3) Were the investigators reliable in determining the formant values when measured by hand?

METHODS

Five men and five women, between the ages of 20-35 with no diagnosed speech and voice disorder, were recruited from the student population in the Department of Communication Sciences and Disorders at West Chester University, West Chester, PA. Following informed consent, the participants wore a head-mounted microphone (AKG, Northridge CA, with a frequency response of 20 – 20,000 Hz and sensitivity at 1000 Hz of 7 mV/Pa) three inches from the mouth and produced sustained phonation of each of the four point vowels (/i/, /μ/, /æ/, /ɑ/) for a duration of 5 seconds at a comfortable pitch and loudness onto a Dell desktop computer. The vowel productions were trimmed to analyze the mid-portion of the vowels (i.e., three seconds of the vowel) to eliminate the effects of voice onset and offset. It was important to represent the experience of a typical clinical SLP who is not trained in formant bandwidth length and sampling rate analysis techniques; therefore, the default settings of formant bandwidth length and sampling rate were used. The formant bandwidth length for analysis was 234 Hz for CSL, 260 Hz for Praat, and 250 Hz for WaveSurfer. The sampling rate frequency was 11,025 for CSL, 44,000 for Praat,

	/i/	/μ/	/æ/	/ɑ/
Five Men	Mean F_o=133.29 Hz SD=46.93	Mean F_o=141.11 Hz SD=43.46	Mean F_o=128.79 Hz SD=39.42	Mean F_o=125.46 Hz SD=43.44
Five Women	Mean F_o=191.53 Hz SD=34.96	Mean F_o=191.91 Hz SD=38.27	Mean F_o=183.04 Hz SD=39.09	Mean F_o=180.64 Hz SD=35.74

Table 1. The mean fundamental frequency (F_o) and standard deviations (SD) of the four point vowels (/i/, /μ/, /æ/, /ɑ/) for the five men and five women participants.

and 16,000 for WaveSurfer. Table 1 contains the average fundamental frequencies and standard deviations for each of the vowels for the five men and five women participants. Two investigators independently analyzed the data for the first three formant frequencies using CSL, Praat, and WaveSurfer. The formant values were assessed for agreement across software programs as well as for inter- and intra-rater reliability within and across investigators.

Analysis Methods

Two methods of capturing the formant frequency values were used to assess the reliability of each method across the software programs. The methods involved: 1) measuring the formants by hand at 2.5 seconds of the vowel and 2) the average formant value reported by the software program for the entire three seconds of the vowel. For calculating the formants by hand, the investigator located 2.5 seconds of the vowel production, used a formant plot to locate the mid-portion of the formant, and then recorded the frequency value of the formant. This procedure was replicated by two investigators for each of the first three formants across all the point vowels in CSL, Praat, and WaveSurfer. Inter- and intra-rater reliability within and across investigator of the analysis method by hand was assessed on 10% of the /ɑ/ vowel productions. For the average value of the formant reported by the software program, the investigator opened the trimmed vowel production (i.e., mid-portion of the vowel three seconds) in the software program and enabled the program to report the average of the formant for the three second duration. This was completed for the first three formant frequencies. Inter- and intra-rater reliability testing for the investigators was not needed because the software program completed the analysis; therefore, eliminating human error. Two of the software programs allowed for averaging of the formant, one program did not. WaveSurfer was not used in the second analysis method. CSL and Praat were compared.

RESULTS

The results will be presented below according to each of the three experimental questions.

Experimental Question One: Will CSL, Praat, and WaveSurfer agree in formant frequency analysis of the four point vowels when measured by hand at 2.5 seconds? A repeated measures analysis of variance revealed significant main effects for software (p=.014), vowel (p=.000), and formant (p=.000). The main effect for investigator was not significant (p=.303). The results indicated that a difference did exist across the three software programs for analysis of the first three formant frequencies. Specifically, the differences were seen between CSL and WaveSurfer (p=.05); therefore, CSL and WaveSurfer produced different formant frequency values when measured by hand at 2.5 seconds. Results related to vowels and formants suggested that all vowels and formant values were different from one another as evidenced by the significant p-values for vowel and formant. This is not surprising given that all participants produced four distinct vowels with increasing formant values from formant one to formant three. Related to investigator, results demonstrated that the two investigators were consistent in measuring the first formant frequencies as evidenced by the non-significant p-value for investigator.

Experimental Question Two: Will the means of the formant values for the four point vowels agree when measured by hand at 2.5 seconds and when averaged by the programs across CSL and Praat? WaveSurfer could not be used in this analysis because the program does not determine an average of the formants across the vowel. A repeated analysis of variance revealed significant main effects for software (p=.000), means (p=.000), vowel (p=.000), and formant (p=.000). Main effect of investigator was not significant (p=.677). The results indicated that a difference did exist across the two software programs for analysis of the first three formant frequencies; therefore, CSL and Praat produced different formant frequency values. A significant two-way interaction was seen for software x mean (p=.000), which suggested that the software programs did not agree across the two different analysis methods. Formant values measured by hand versus formant values averaged across the vowel by the programs did not agree. Results related to vowels and formants were consistent with the results for experimental question one that all vowels and formant values were different from one another. Related to investigator the results were consistent with experimental question one that the two investigators agreed in their measurement of the first formant frequencies.

Experimental Question Three: Were the investigators reliable in determining the formant values when measured by hand at 2.5 seconds? The vowel /ɑ/ was selected for reliability testing across the investigators. Across all paired samples tests, the results were highly correlated with a 1.0 correlation and p=.000 significance. The investigators demonstrated highly correlated

inter-rater reliability. The results from the ANOVAs above suggest good intra-rater reliability as determined by the non-significant findings involving investigator.

DISCUSSION

The results suggested that when measured by hand, CSL and WaveSurfer did not agree in formant value analysis. CSL versus Praat did agree in formant analysis and Praat versus WaveSurfer did agree in formant analysis; therefore, Praat demonstrated the best agreement among the three software programs when the formants were measured by hand. The findings suggest that comparisons may not be appropriate between CSL and WaveSurfer because the formant analysis differs when measured by hand at a specific time interval. In addition, when analysis method (i.e., by hand versus average across vowel reported by program) was compared with software (i.e., CSL versus Praat), the formant value analysis did not agree. WaveSurfer does not have the option of averaging the formant value across a certain length of time; therefore, only CSL and Praat were compared. Based on the findings, it may be more reliable and valid to determine formant frequency values by hand rather than reporting the average across the vowel as the agreement was better between the software programs, except between CSL and WaveSurfer. The differences in findings could not be attributed to investigator error due to high inter- and intra-rater reliability of measuring the formant values.

Based on the results, implications for practice are presented below. Praat had the best agreement among the three software programs when measured by hand. Due to this finding, results of patient performance should at least be reported using Praat for analysis of formant frequencies to facilitate consistent comparisons across patients and SLPs. Perhaps relying on the average formant value reported by the program should be abandoned because CSL and Praat did not agree in this analysis technique. The SLP may consider measuring the formants at a specific time interval to achieve better agreement among programs; thereby, reporting consistent results across patient and clinician. A major benefit of Praat and WaveSurfer is that they are free and accessible on the internet. Praat and WaveSurfer can also be downloaded to a patient's computer and used at home for practice of various therapy goals. CSL is available for purchase, but it is expensive; therefore, patients cannot use the CSL at home for carry-over work outside the clinic setting. Overall, in choosing the best software program to use, the SLP first must determine the rationale behind its use. For example, if the software program will be used for reporting of results in a formal report, then Praat should be considered due to its high agreement among the three programs when the formants were measured by hand. If the SLP wants to use the software program as a method in targeting a speech and voice production goal, then all of the three programs are appropriate to consider. An SLP may not have access to CSL, then WaveSurfer and Praat would be the best options because they are free and available on the internet. If the SLP wants to use the software programs as a carry-over method for patients to use at home for practice of various therapy targets, then WaveSufer and Praat would be the best options. Of the two free programs available online, WaveSurfer was the easier to use.

ACKNOWLEDGMENTS

An abbreviated version of the findings was presented at the Voice Foundation's 41st Symposium: Care of the Professional Voice in Philadelphia, Pennsylvania, June 2012 and the American Speech Language Hearing Association (ASHA) convention in Atlanta, Georgia, November 2012.

CORRESPONDENCE

egrillo@wcupa.edu

REFERENCES

American Speech-Language-Hearing Association. (2007). Scope of Practice in Speech-Language Pathology [Scope of Practice]. Available from www.asha.org/policy.

Fant, G. (1960). *Acoustic theory of speech production* (2nd edition). The Hague: Mouton.

Ladefoged, P. (1996). *Elements of acoustic phonetics*. Chicago: University of Chicago Press.

Ladefoged, P. (2006). *A course in phonetics* (4th edition). Sydney, Australia: Thomson-Wadsworth.

Ladefoged, P., Harshman, R., Goldstein, L., & Rice, L. (1978). Generating vocal tract shapes from formant frequencies. *Journal of the Acoustical Society of America*, 64, 1027-1035.

Smits, I., Ceuppens, P., & De Bodt, M.S. (2005). A comparative study of acoustic voice measurements by means of Dr. Speech and Computerized Speech Lab. *Journal of Voice*, 19(2), 187-196.

AUTHOR INFORMATION

Dr. Elizabeth Grillo (Ph.D. The University of Pittsburgh) is an Associate Professor in Communication Sciences and Disorders at West Chester University and an Adjunct Research Instructor at the Jefferson Center for Voice and Swallowing in Philadelphia, PA. One aspect of her research involves acoustic and aerodynamic investigations of speech and voice. Ralph Magliano and Jodi DiSario were M.A. students at West Chester University at the time of this research.

JOURNAL
CALL FOR SUBMISSIONS
2014 EDITION
(DEADLINE: 15 SEPTEMBER 2014)

PURPOSE:

The PSHA JOURNAL is an annual electronic peer-reviewed publication of the Pennsylvania Speech-Language-Hearing Association. The revival of the PSHA Journal complements the aims of PSHA as (a) a means to broaden the nature of the services PSHA provides to PSHA members, (b) an outlet to showcase information about clinical education, clinical practices, and responses to professional issues, particularly as relevant to Pennsylvania, (c) a means to distribute scholarship in its diverse forms and to facilitate access to appropriate publication outlets, and (d) a tool to advance the information that will assist PSHA members in their quest to best address the clinical needs of individuals with communication disorders across professional roles and responsibilities.

SUBMISSIONS:

The PSHA JOURNAL accepts manuscripts for review that reflect diverse scholarly work. Manuscripts that reflect these scholarly orientations are welcome:

Basic Scholarship: *Basic Scholarship* includes those articles that expand our information base with respect to normal and non-normal communication processes, as well as normal and non-normal processes in the communication mechanism. Articles that illustrate this variation of scholarship are data-based studies (with quantitative and/or qualitative data) that, while not necessarily experimental in nature, may seek to confirm or disconfirm specific theoretical assertions.

Conceptual Scholarship: *Conceptual Scholarship* includes those articles that present and/or expand upon theoretical considerations, as well as those articles that conceptualize "older" ideas and/or data into "newer" perspectives. Articles that illustrate this variation of scholarship include theoretical discussions and debates, state of the art reviews of literature, meta-analyses of extensive corpora of data, and/or discussions and debates of salient professional issues and directions.

Applied Scholarship: *Applied Scholarship* includes those articles that present and address clinical and professional issues, typically with data-based studies as their focus. Data may be quantitative and/or qualitative in nature, and studies may be experimental and/or descriptive in nature. The results of applied scholarship often have clear implications for and applications to current issues. This scholarship may also consist of series of studies that build upon a theme.

Instructional Scholarship: *Instructional Scholarship* includes a diverse assortment of articles that aim to improve our approaches to and materials and methods for instruction. Because our disciplines have clinical practice at their heart, instructional scholarship also includes articles that aim to improve clinical service delivery across work sites, work duties, and clinical populations. Instructional scholarship also includes reports of innovative clinical education and professional development models.

In addition to these variations of scholarship, the PSHA Journal welcomes for review any other content the PSHA Executive Board deems important to advance our disciplines and/or serve our membership.

AUTHOR INSTRUCTIONS:

Authors: The PSHA Journal welcomes submissions from both PSHA members and PSHA non-members.

Author Information: In the cover letter to be submitted with the manuscript, please provide the names, academic degrees, CCCs and other professional credentials, professional titles, and professional affiliations for the manuscript authors. Each author is welcome to submit a professional photo to be published with his/her accepted manuscripts to further familiarize the readers with his/her identity. Please submit a 100- to 150-word biographical statement for each author that summarizes his/her professional contributions. Also, please provide the phone number, postal mail address, and electronic mail address for each author, as well as the PSHA membership status for each author.

Authorship Information: In the cover letter, please include verification that, in the case of multiple authors for a manuscript, the authors

have reviewed and have consented to the order of authorship for the manuscript. Please note that the order of authorship for the manuscript implies that the senior author(s) was/were the primary contributor(s) and that the junior author(s) was/were the secondary contributor(s).

Author Rights: In the cover letter, please include verification that the authors have consented to the publication (contingent upon positive review and subsequent acceptance) of their manuscript in the PSHA Journal. At this time, articles included in the PSHA Journal will be copyright in the name(s) of the author(s). Authors should contact PSHA at any future time for permission to reproduce portions of their articles in other venues.

Human Subjects: When a manuscript includes data (quantitative and/or qualitative) collected from human subjects, the authors of the manuscript should provide verification (in the form of a duplicate approval form from the relevant Institutional Research Board) of the protection of these subjects in each phase of the research process.

Manuscripts: Manuscripts should be the original work of the named authors. Plagiarism in any form is automatic ground for non-acceptance of a manuscript. Manuscripts submitted to the PSHA Journal should not be previously published in any form in any other venue.

Manuscript Preparation: Authors should submit manuscripts in electronic form. Authors should use Microsoft Word for the transmission of their manuscripts, APA style for the format of their manuscripts, and 12 point Times New Roman font for the text. At this time, the minimum word count for manuscripts is 2,500 words, and the maximum word count for manuscripts is 10,000 words. Authors who wish to exceed the maximum word count for a manuscript should consult with the Editor before submission. Each manuscript should contain an abstract of 100-150 words. Authors should indicate 3-4 Key Words to describe the primary content of the manuscript. A reference list of all works cited within the text/appendices should be included after the main text. Figures/Tables/Images/Other should be appended to the main text with their approximate locations indicated in the text. Images that are photos should be submitted in jpeg format with a minimum of 600 pixels per square inch. The manuscript should not contain footnotes or endnotes. The file should have editorial tracking utilities disabled. Authors should acknowledge any funding sources for support of their work, as well as any individuals whose contributions enhanced the work.

Manuscript Submission: Authors should submit their manuscripts on or before 15 September 2014 to Cheryl Gunter, Ph.D. CCC-SLP, PSHA Vice-President for Publications/Editor, at cgunter@wcupa.edu. For further information, please contact the Editor at 610.436.2115 or at this e-mail address. Information about the manuscript review process will be available to authors at the time of their manuscript submission.