Cross-linguistic generalization in the treatment of two sequential Spanish–English bilingual children with speech sound disorders

CHRISTINA GILDERSLEEVE-NEUMANN1 & BRIAN A. GOLDSFTEIN2

1Portland State University, Portland, OR, USA, and 2La Salle University, Philadelphia, PA, USA

Abstract
The effect of bilingual service delivery on treatment of speech sound disorders (SSDs) in bilingual children is largely unknown. Bilingual children with SSDs are typically provided intervention in only one language, although research suggests dual-language instruction for language disorders is best practice for bilinguals. This study examined cross-linguistic generalization of bilingual intervention in treatment of two 5-year-old sequential bilingual boys with SSDs (one with Childhood Apraxia of Speech), hypothesizing that selecting and treating targets in both languages would result in significant overall change in their English and Spanish speech systems. A multiple baseline across behaviours design was used to measure treatment effectiveness for two targets per child. Children received treatment 2–3 times per week for 8 weeks and in Spanish for at least 2 of every 3 days. Ongoing treatment performance was measured in probes in both languages; overall speech skills were compared pre- and post-treatment. Both children's speech improved in both languages with similar magnitude; there was improvement in some non-treated errors. Thus, treating both languages had an overall positive effect on these bilingual children's speech. Future bilingual intervention research should explore alternating treatments designs, efficiency of monolingual vs bilingual treatment, different language and bilingual backgrounds, and between-group comparisons.

Keywords: Intervention, speech sound disorder, bilingualism.

Introduction
Speech sound disorders (SSDs) affect ~680,000 children in the US, including ~10% of pre-school and school-age children (Gierut, 1998). For children with SSDs (whether phonological, articulatory, and/or motor planning in origin), receiving treatment is crucial for later academic success. Between 50–70% of school-age children with a SSD exhibit academic difficulty and often require other remedial services (Gierut, 1998). Children's quality-of-life is affected by having a speech sound disorder (Markham & Dean, 2006; Markham, van Laar, Gibbard, & Dean, 2009), with the frustration from not being understood often resulting in negative social effects (McCormack, McLeod, McAllister, & Harrison, 2010).

A growing number of children in the US are in bilingual language environments and ~12% of the US population speaks Spanish (Shin & Kominski, 2010). For bilingual children with SSD, the challenges for receiving adequate treatment are considerable. Although 92% of US school-based speech-language pathologists (SLPs) indicate having children with SSDs on their caseloads, they are more likely to provide treatment only in English, not in the child's other language, given that only ~5.2% (6300 of 120,744) of ASHA-certified SLPs report being bilingual. Of these 6300 bilingual SLPs, ~2700 speak Spanish (ASHA, 2010).

Although the regular delivery of treatment in English to bilingual children with SSD is convenient, it might not be warranted by the evidence. Research shows that Spanish–English bilingual children succeed academically at rates higher than children with limited English proficiency or children from Latino backgrounds who do not speak Spanish (Rumberger & Larson, 1998). Research indicates that academic instruction in the non-English (i.e., home) language helps rather than hinders the acquisition of English (e.g., Campos, 1995; Cobo-Lewis, Eilers, Pearson, & Umbel, 2002; Kohnert, 2013), and, in fact, leads to higher academic achievement in English (e.g., Thomas & Collier, 2003). For treatment of child language disorders, current best practices argue for providing intervention in the home language as well as in English (Kohnert, 2013; Perozzi & Sanchez, 1992; Restrepo, Morgan, & Thompson, 2013).

It is not known whether intervention in only English is an effective method of delivering speech
therapy to bilingual children with SSDs. Although it is accepted that bilingual children have differentiated linguistic systems by age 2;0 (Keshavarz & Ingram, 2002), the two speech sound systems in bilingual children are likely inter-linked, which would allow intervention effects in one language to generalize to the other language (e.g., Fabiano-Smith & Goldstein, 2010; Paradis, 2001). However, the degree of such generalization between speech systems for bilingual children with SSD after receiving treatment for their disorder is unexplored. This study examines cross-linguistic generalization of speech skills as well as the effect of bilingual intervention on two children’s bilingual speech systems.

Theoretical framework and its application to bilinguals

Dynamic systems theory (DST) provides the theoretical foundation for this investigation of bilingual intervention for SSD, although the theory is not being tested directly. DST explores how seemingly autonomous components are inter-connected in developmental stages of complex biological systems (Thelen & Smith, 1995; Thelen, Kelso, & Fogel, 1987), including speech (Thelen, 1991). The complexity of speech development from a DST perspective guides our intervention approach, which utilizes articulatory, phonological, and metalinguistic factors. DST also guides our hypothesis that the two languages of a bilingual child are linked in development and that intervention that addresses both languages will have cross-linguistic effects on speech development. We propose that bilingual intervention can also be guided by DST by selecting speech targets that affect both languages and by providing dual-language intervention.

DST has been explored in treating SSDs (Rvachew & Bernhardt, 2010). In a randomized control trial, Rvachew and Bernhardt treated 48 children with SSDs, randomly assigning them to two groups based on the complexity of their treatment targets. Results indicated that the children assigned to the group focusing on less complex targets made more progress than those in the group focusing on more complex targets. Specifically, a more complex form, /tʃ/, emerged from less complex characteristics of earlier speech development.

In development, dynamic interaction between inter-linked components is expected to result in skill progression. However, DST also accounts for U-shaped change: surface-level regressions concurrent with underlying progress in the complexity of higher-level abilities (Gershkoff-Stowe & Thelen, 2004). This U-shaped change has been noted in the language skills of developing bilinguals (Kohnert, 2013). In bilinguals, assumptions of DST suggest underlying, dynamic interactions between languages would result in cross-linguistic generalization during development. Such cross-linguistic generalization would allow bilingual children to transfer phonological and articulatory knowledge from one language to increase accuracy of productions in the other. This cross-linguistic generalization helps explain why a bilingual child, who does not speak either of their two languages as much as monolingual children do, develops speech at about the same rate as monolinguals (e.g., Fabiano-Smith & Goldstein, 2010; Gildersleeve-Neumann & Wright, 2010; Gildersleeve-Neumann, Peña, Davis, & Kester, 2009). Our study examines cross-linguistic generalization in bilingual treatment for two Spanish–English bilingual children to explore bilingual treatment effectiveness and dynamic interactions across languages.

Evidence for intra-linguistic generalization in SSD treatment in monolingual children

The expectation for intra-linguistic generalization is a basic assumption of speech-language therapy and many intervention studies on monolingual children with SSD have systematically investigated the rate and types of intra-language generalization. Tyler, Edwards, and Saxman (1987) examined two phonological process-based approaches with four children ranging in age from 3;1–5;1. Treatment lasted for 6–8 weeks and included 12–16 total sessions. A generalization probe was created to track progress on treated and untreated sounds. The intervention resulted in an increase in performance on treated and untreated sounds. Performance on treated sounds was greater than that for untreated sounds. As expected, cross-class generalization did occur for these participants, as sounds across classes were chosen as intervention targets.

Gierut and colleagues have evaluated a number of factors related to intra-linguistic generalization in treatment studies with monolingual, English-speaking children with SSD (Gierut & Champion, 2001; Gierut, Elbert, & Dinnensen, 1987; Gierut, Morrisette, Hughes, & Rowland, 1996). Overall, these studies suggest that targeting more complex aspects of the speech system (e.g., sounds excluded from the inventory, later-developing sounds, non-stimulable sounds, more marked sounds) results in greater system-wide changes than does targeting less complex aspects of the speech system (sounds included in the inventory, early-developing sounds, stimulable sounds, less marked sounds). However, in a randomized design treatment study of 48 children with moderate-to-severe SSD, Rvachew and Nowak (2001) found that children who received intervention focusing on early-developing sounds showed greater improvement than those participants whose targets were later-developing. Moreover, greater intra-linguistic generalization did not occur for those receiving intervention on later-developing sounds. Rvachew and Bernhardt’s analysis of treatment...
response suggests that “new complex behaviours can emerge when intervention serves to stabilize sub-component skills” as predicted by DST (Rvachew & Bernhardt, 2010, p. 37). From a bilingual perspective, these stable sub-components could be interpreted as having greater inter-language associations and being more salient than those having lesser inter-language associations.

Childhood apraxia of speech (CAS) is an SSD resulting primarily from motor planning difficulties (ASHA, 2007). Studies of monolingual children with CAS suggest intense, frequent, and long-term treatment that emphasizes the motor planning of speech, and that generalization of treatment targets to new phonological properties is limited (ASHA, 2007; Ballard, Robin, McCabe, & McDonald, 2010; Strand, Stoeckel, & Baas, 2006). Strand and Debertine (2000) used a single-subject design with a 5-year-old with CAS, training a limited set of functional phrases in short intense sessions 4 days per week. Target phrases improved, but no generalization was shown. In a single-subject design study with four children with CAS, Strand et al. (2006) trained a small number of phrases 2-times per day, 5 days per week. After 6 weeks, three of the four children demonstrated improvement and limited generalization effects. Also using a single-subject design, Edeal and Gildersleeve-Neumann (2011) treated two children with CAS between 2–3-times per week for 40-minute sessions. High intensity treatment (100 + productions in 20 minutes) resulted in the most generalization of speech skills. Research has not demonstrated the best treatment strategies for generalization across languages in bilingual children with CAS, although intense articulatory treatment, likely focusing on early-developing sounds to build a stable foundation for further speech development, builds on the findings from monolingual research on children with CAS.

Cross-linguistic generalization in treatment for SSD in bilingual children

Only a few studies have examined cross-linguistic generalization in the treatment of bilingual children with SSD; no studies have examined cross-linguistic generalization in children with CAS. Holm, Ozanne, and Dodd (1997) treated a 5-year-old Cantonese–English speaking boy who heard and spoke Cantonese at home until the age of 3½, when he entered school. The child was treated in English-only, in two phases. The first phase targeted distorted /s/, and the second phase targeted cluster reduction. Cross-linguistic generalization was mixed. Findings indicated that treatment on /s/ resulted in changes on /s/ in both languages; however, the treatment of cluster reduction only improved cluster production in English and not in Cantonese. These results are confounded by the commonly-accepted belief that Cantonese does not contain clusters in its phonology (So, 2007).

Holm and Dodd (2001), using a pre-/post-case study design, examined cross-linguistic generalization in one Cantonese–English and one Punjabi–English speaking child, both first exposed to English around age 3. In the Cantonese–English child, treatment initially consisted of articulation therapy in English, focusing on distortion of interdental /s/. The authors hypothesized that generalization would occur to /s/ in Cantonese, given that the error was not language-specific. Results indicated that, after 7 weeks of intervention, the child improved in English on the target sound and sounds not directly targeted: /ʃ/ and clusters. Generalization occurred in Cantonese as well; the child produced /s/, /ts/, /tsʰ/ in Cantonese with 70% accuracy (percentage accuracy of those segments prior to intervention was not indicated). Subsequently, the child received eight weekly sessions of phonological therapy. Cluster reduction and gliding in English decreased after intervention. The authors noted that generalization across languages did not occur, although they did not provide post-intervention data in Cantonese supporting this position. With the second child (Punjabi–English), Holm and Dodd used a core vocabulary intervention approach to carry out English-only treatment. This approach emphasized sound production and phonological awareness skills. Results indicated increased consonant accuracy in both English and Punjabi immediately after and 2 weeks post-treatment.

Ray (2002) used a cognitive-linguistic treatment approach with a 5-year-old trilingual (Hindi, Gujarati, and English) boy with a mild SSD. The child, born in the US, was exposed to Hindi and Gujarati from birth. At age 4, he was “formally introduced” to English at school, but he had been exposed to it prior to that time. Minimal contrast therapy in English was used to target multiple error patterns (final consonant deletion, gliding, cluster reduction). Results indicated that treatment was effective in decreasing the percentage-of-occurrence of error patterns and increasing consonant accuracy in all three languages, and improving the child’s percentile rank on a standardized English articulation assessment. Although phonological skills were examined in all languages, the effect of clinician and home practice was not accounted for, nor was it clear if and how generalization took place across the child’s languages.

Limitations of previous research

Although the few existing studies focusing on remediating SSD in bilingual children indicate that treatment in English promotes generalization, they were not designed to examine the therapeutic benefits of cross-linguistic generalization. In all cases, treatment took place only in English, even though the children were bilingual or trilingual. As noted for other areas of language, improvement in skills
in the home (non-English) language resulted in improvement of skills in English. It remains unclear if these generalization effects were only unidirectional, leaving an open question as to whether generalization effects occur when treatment is in the home language and English. We also do not fully understand how cross-linguistic generalization operates.

Findings from multilingual treatment studies to date have limited generalizability for a variety of reasons. The case studies have not utilized experimental designs or control conditions, making it difficult to determine if the treatments themselves resulted in the changes and whether the findings generalize to other bilingual children. None of the studies reported measures of treatment fidelity. Few details on the specific treatment approaches have been specified, target choices are often not described or justified, and the definition of generalization employed has not been provided. In addition, it is unknown whether there would be greater overall, or more rapid, improvement if the treatment had first occurred in the home or community language, or bilingually, or if it had been based on the choice of treatment targets. Also, effects of other variables known to influence intra- and cross-linguistic generalization in SSD treatment (e.g., language proficiency, SSD severity) are often not controlled. Changes in the children’s phonological system not targeted as part of intervention are not typically measured. Finally, no studies of cross-linguistic generalization in bilingual children with CAS have been published.

In summary, results from treatment studies of SSDs on monolingual and bilingual children suggest that intervention highlighting language-specific interactions during speech intervention can yield intra- and cross-linguistic generalization of that speech information. Results also indicate that interactions focusing on intra- and/or cross-linguistic generalization can bring about changes in the whole speech system because of that input during intervention. These general findings have proven to be robust in monolingual English-speaking children. However, the exact nature of the importance of these variables in SSD remediation in bilingual children has not been measured. We hypothesize that selecting targets with strong interaction capacity (phonological properties whose errors affect both languages at relatively high error rates) and treating targets in both languages will result in cross-linguistic generalization, as measured by significant increases in phoneme accuracy and phonetic complexity and significant decreases in the percentage-of-occurrence of error patterns in both English and Spanish.

**Method**

**Participants**

The children in this study were two boys: Carlos and Rubén (pseudonyms). Both boys were diagnosed with SSDs in their educational settings prior to study onset; Rubén also has CAS. They were born in the US in predominantly Spanish home language environments. Both boys were referred to the university clinic by their school-based SLPs. Both boys had received their initial assessment at their educational setting prior to this treatment study and had received intervention for ~6 months in English. The boys were generally healthy.

Carlos was 5;8 at study onset. Carlos had just completed a full-day kindergarten in a Spanish–English bilingual immersion school. He lived with his parents, 8-year-old brother, and 1-year-old sister. Carlos’ parents were originally from Mexico and spoke and understood little English; all spoken and written communication with Carlos’ parents was conducted in Spanish, based on parental preference. Carlos was first exposed to English regularly 2 years prior to this study. Carlos occasionally spoke limited English with his older brother when they were with English-speaking friends; in all other settings Carlos spoke Spanish. He had been diagnosed with a moderate SSD ~9 months prior to study onset.

Rubén was 5;6 at the onset of this study. He lived with his parents who were originally from Mexico, and his 1-year-old sister. Rubén’s family spoke no English at home, nor did his younger sister. All spoken and written communication with Rubén’s parents was in Spanish, by parental preference. Rubén’s father understood and spoke English moderately well; his mother reported not speaking or understanding English. Rubén had first been exposed to English regularly 1 year prior to the study when he entered an English-only Head Start classroom for 4 hours per day, 3 days per week. Rubén had been diagnosed with a severe SSD at 3;6; his school clinician also suspected CAS, which was confirmed during this study. Rubén had received speech therapy in English once per week in his classroom.

**Design and procedures**

As part of this treatment efficacy study, three outcome variables were measured pre- and post-intervention: phoneme accuracy; phonetic complexity; and phonological error patterns. Two treatment targets, Target 1 (TRG1) and Target 2 (TRG2), were selected for each child based on individual assessment results. Carlos and Rubén received treatment 2–3-times per week for 8 weeks. Treatment was provided in Spanish for at least 2 out of every 3 days; English was provided on the remaining days. To determine treatment effects during the intervention, we used a multiple-probe across behaviours design, a type of single-subject experimental design. Baseline measurements (the control conditions) for multiple behaviours allow for description of pre-treatment levels and predictions for future levels of performance with the intervention (Kazdin, 2011). A strength of the single-subject design is that the effects of treatment
can be measured in an individual child’s intervention (Horner, Carre, Halle, McGee, Odom, & Wolery, 2005), lessening the threats to validity for individuals that group treatment research in controlled settings have.

We established baseline for each child’s treatment targets in both Spanish and English. Treatment for TRG1 began first while baseline data continued to be collected for TRG2. During this first treatment phase, non-targeted language activities were completed during treatment time set aside for TRG2. Eventually, both TRG1 and TRG2 were treated for both children (and no additional language activities were conducted). The targets selected were different for each child. The criterion for target selection was existence of the error in both languages and error pattern rate, with developmental appropriateness considered in the selection of errors with high rates of occurrence (Rvachew & Nowak, 2001). Probe data were collected for the two treated targets. Stimuli at the appropriate level (sound, word, or phrase) were presented to each child with a request to repeat the stimuli. The probes were randomly selected for collection at the start of each session.

Treatment sessions were 50 minutes long, divided as follows: 5-minute probes of treated target behaviours, 20 minutes for each treatment target (randomly ordered), 5 minutes for metalinguistic treatment generalization across languages: discussion and application of treatment information to targets in the non-treated language. In Phase 1, only TRG1 was treated; time allotted for TRG2 was spent on non-goal receptive language activities (e.g., sequencing, colour, and attribute identifications). Phase 1 was complete when the child achieved a pre-determined probe accuracy criterion for the level they were at with TRG1. During Phase 2, both TRG1 and TRG2 were treated. Treatment involved articulatory and phonological strategies for both children; for the child with CAS, the cueing strategies effective in treatment for motor planning difficulties were added to facilitate productions (e.g., Edeal & Gildersleeve-Neumann, 2011; Strand et al., 2006). All assessment and treatment procedures were administered to each child by a graduate student clinician under the supervision of the first author. The student clinician was trained in the intervention protocol. Procedures included four consecutive phases of direct individual interactions with participants: eligibility assessment (analysis of which determined treatment targets), baseline testing, treatment, and post-treatment speech testing. Each phase is described below.

**Phase 1: Eligibility assessment.** To determine appropriateness for a speech production study, we conducted a full assessment of each boy’s communication system prior to treatment in the clinic setting. We gathered detailed case history information for each child from parent surveys, evaluating developmental history, language use, and parent concerns related to speech. The speech development questions on the unpublished parent survey allowed Likert scale comparisons between responses. Previous research has found strong correlations between this parent survey and percentage of consonants correct (PCC) in Spanish (Stertzbach & Gildersleeve-Neumann, 2005) and English (Powers, 2010).

Both Rubén and Carlos passed an oral peripheral examination and pure-tone hearing screenings at 20 dBA HL in each ear at 1000, 2000, and 4000 Hz. Informal evaluation confirmed age-typical cognition, voice, and fluency. Rubén was shy, which initially affected appropriate social communication; however, both boys were determined to have typical social communication skills. Language skills were evaluated using formal and informal assessment measures. Rubén and Carlos were administered the Spanish Clinical Evaluation of Language Fundamentals-Preschool 2 (Wiig, Secord, & Semel, 2009). In keeping with best practices for use of standardized assessments for bilingual children, scores are not reported because of the test’s low validity, specificity, and sensitivity for bilingual children. Nevertheless, the performance of both boys fell within the typical confidence interval for monolingual Spanish speakers. These findings were corroborated with observations during conversation and narrative tasks in the test administration, suggesting that Rubén’s and Carlos’ receptive language skills were within normal limits.

Speech production was assessed through standardized but not norm-referenced assessment measures. These measures included single-word and sentence-level articulation and phonology tasks in Spanish and English, using the Phonological and Articulatory Bilingual Assessment (PABA) (Gildersleeve-Neumann, 2014). The PABA is a research protocol with English and Spanish components that are administered on different days. It contains single word and sentence components with stimuli that represent the children’s vocabulary, includes 1-, 2-, 3-, and 4-syllable words, all language-appropriate phonemes (except English /ɣ/ because of its low frequency in children’s vocabulary), and within- and across-syllable consonant sequences frequent in each language. The single word portion of the PABA is a picture identification task. There are 136 English words and 115 Spanish words. If children do not identify the pictures spontaneously, the clinician produces the picture name in a delayed imitation format; if the child does not produce the word in delayed imitation, direct imitation is utilized. The imitated sentence subcomponents have three trial sentences and 18 sentences with pictures reflecting all key words. If the child does not produce every word in the imitated sentence, the non-produced words are removed from the target sentence for analysis. All speech samples were digitally recorded in audio and video.
formats for later transcription. Speech skills were evaluated in Spanish and English separately.

**Phase 2: Pre-treatment target selection and baseline criterion.** Results from the speech eligibility assessment were used for four fundamental purposes: (1) to select individual treatment targets for each child, (2) to provide the first of minimally three target measurements to serve as a baseline for treatment efficacy data collection, (3) to provide pre- and post-measurements for three dependent variables in Spanish and English (PCC, percentage of vowels correct (PVC), and phonetic complexity), and (4) to provide a pre- and post-treatment measure for comparing to treatment performance.

**TARGET SELECTION**

For each child, two targets were addressed. Criteria for target selection were the existence of the error in both languages and the error rate (emphasizing higher rates), with developmental appropriateness considered in selection (Rvachew & Bernhardt, 2010; Rvachew & Nowak, 2001). Targets were addressed in both languages. Probe stimuli were individualized for each child and target. Probe stimuli for a particular target changed if short-term objectives were achieved.

**Phase 3: Treatment**

**INTENSITY**

Treatment intensity is a critical variable in successful speech therapy that affects outcomes. Intense treatment more likely results in the desired change (Duhon, Mesmer, Atkins, Greguson, & Olinger, 2009). Research suggests that greater frequency of production targets results in increased accuracy in probe data as well as generalization to non-targeted utterances (Edeal & Gildersleeve-Neumann, 2011).

There are many ways that intensity is interpreted in the treatment literature (Maas, Robin, Hula, Freedman, Wulf, Ballard, et al., 2008; Warren, Fey, & Yoder, 2007). Treatment intensity differences include individual vs group intervention (Duhon et al., 2009), total number of treatment hours (e.g., Lovaas, 1987), number of elicited target productions during therapy (Edeal & Gildersleeve-Neumann, 2011), and number of hours per week of treatment (Robey, 1998). As described below, we controlled treatment duration and frequency, number of speech productions per treatment target, and treatment form.

**TREATMENT DESIGN**

The organization of procedures in a single 50-minute session was as follows:

1. Collect probe data for treatment goals. During each session, sets of 10 probe stimuli in Spanish and/or English were collected. Probe sets for each target were not collected in every session. Probe data collection took ~5 minutes.

2. Treatment for Target 1. Twenty minutes were spent on TRG1. This period began with a brief review of the target and then drill play activities to elicit 40–50 attempts of accurate production targets in functional words and phrases as appropriate for treatment level.

3. Treatment for Target 2. Twenty minutes were spent on TRG2. The same activity types and frequency of production goals used for TRG1 were used for TRG2.

4. Carry-over discussion of treatment goals to language not treated. The final 5 minutes of each session focused on the meta-linguistic application of treatment goals to the non-targeted language.

**TREATMENT PHASE 1**

In Phase 1, the first target to reach a stable baseline based on target probe measurement was treated. A baseline was considered stable if there was an unchanging or increasing error rate for the target over three baseline probes. If both treatment targets had similar baseline stability, TRG1 was selected by the first author. In Phase 1 only TRG1 was treated; a non-speech language activity was completed during the session time allotted for TRG2. Criterion for Phase 1 completion was individually determined and was not necessarily probe-related.

**TREATMENT PHASE 2**

Once the initial criterion for TRG1 was reached, a new criterion for TRG1 was set for Phase 2 and treatment for TRG2 began, also with an established probe criterion. TRG1 and TRG2 were each treated for 20 minutes; order of target treatment within a session was randomized. Probes for TRG1 and TRG2 were collected regularly at the beginning of each session to note response to treatment in both languages and to monitor error rates for the untreated target, when applicable.

**APPROACH**

There are several individual treatment procedures that have been shown to be effective in facilitating speech improvements in young children with SSD (Gierut, 2005). Our treatment approach combined features common to successful articulatory and phonological treatment for monolingual children. Core features to this approach were: (a) promoting child’s meta- and perceptual awareness of specific session speech goals; (b) using developmentally-appropriate and engaging practice activities that facilitate drill play conditions, most effective for speech sound.
treatment in pre-school-aged children; (c) including articulatory and phonological components to each treatment session with cueing strategies that have been demonstrated as effective for children with motor planning difficulties (Edeal & Gildersleeve-Neumann, 2011; Strand & Debertine, 2000); and (d) utilizing intensive production practice on targets embedded in functional words and phrases (Elbert, Powell, & Swartzlander, 1991; Skelton, 2004). Our approach assumed that children need extensive practice in each treatment session to be successful with new motor patterns they had not yet produced successfully (Edeal & Gildersleeve-Neumann, 2011; Skelton, 2004).

Although the exact treatment protocol depended on each child’s target stimulability, treatment involved articulatory placement training, as well as cueing and feedback to ensure accurate production in isolation or single syllables. After the child produced the sound or syllable shape accurately, the target was practiced in functional words and phrases in which phonetic context complexity was monitored. A variety of words and phrases of varying lengths were utilized. If the child made errors, utterance complexity was simplified until success was re-achieved; at that point, the level of phonetic complexity was again increased.

Concurrently, discussion of phonological contrasts was utilized to increase phonological knowledge of target behaviours. We utilized minimal pair strategies (e.g., Weiner, 1981) because they have been used frequently in treating SSDs in children when phonological contrasts are neutralized (e.g., Saben & Ingham, 1991; Williams, 2000). We used minimal pairs when possible to contrast a target sound to the children’s productions that were in error. Combining articulatory and phonological strategies has been utilized in previous studies in the treatment of children with SSDs (e.g., Saben & Ingham, 1991; Tyler et al., 1987; Weiner, 1981). As targets were produced with more accuracy, practice in a variety of utterance lengths was used to support generalization (Skelton, 2004). This varied practice was used unless there was a regression in target accuracy; in these cases, blocked practice at a shorter less complex level was conducted until success was achieved again. Frequency of 40–50 productions in functional words and phrases per 20-minute period was targeted and consistently achieved. To assure that length of session components was adhered to, a timer was used. This production frequency goal was monitored through a hand-clicking mechanism employed by the therapist.

TREATMENT FIDELITY

All sessions were videotaped and reviewed by the first author to verify that the session included each of the core treatment features specified above.

Phase 4: Post-treatment speech testing

POST-TREATMENT PROBE MEASURES

The same probe types and administration were used post-treatment to determine if treatment progress was maintained. Single-word and sentence speech testing administered prior to treatment were repeated to determine children’s post-treatment performance on the three dependent variables: phoneme accuracy; phonetic complexity; and error patterns in Spanish and English. Pre-treatment and post-treatment scores were compared for each child to measure change.

Coding, scoring and analysis

PRE- AND POST-TREATMENT SPEECH TESTING

The single words and imitated sentences were recorded digitally and uploaded for computer playback. All single-word and sentence speech samples were phonetically transcribed by trained individuals who were proficient in Spanish and English to ensure accurate transcription of English and non-English phonemes. The transcribed data were entered into Logical International Phonetics Program software (Oller & Delgado, 2000) to analyse and compare Spanish and English sound properties. Comparisons were made to the adult target, taking dialect into account.

Pre- and post-treatment performance for each child were determined as follows in the single word and sentence tasks in both languages, examining (1) phoneme accuracy using PCC (Shriberg, Austin, Lewis, McSweeney, & Wilson, 1997) and PVC (Shriberg, 1993) for English and Spanish, (2) average utterance complexity in both languages, using Index of Phonetic Complexity (IPC) (Jakielski, 1998), and (3) phonological error pattern percentage-of-occurrence. We determined treatment effect size measuring the standard mean difference for all data points (SMD\textsubscript{ALL}), determining the mean difference in probe data points for baseline and treatment periods and dividing by the standard deviation, for each child’s TRG1 and TRG2 single word probe data and for the combined total. The use of SMD\textsubscript{ALL} effect size allows for inclusion of all data points. Effect sizes were interpreted using Cohen’s $d$ as follows: small ($.02–.049)$, medium ($.50–.79)$, and large ($.80$ or greater) (Cohen, 1988).

TRANSCRIPTION RELIABILITY

Inter-judge and intra-judge reliability for broad and narrow transcription (Shriberg & Lof, 1991) was completed on the pre- and post-assessment measures and the probe words. All productions were re-transcribed by the first author and were subjected to established methods for determining inter-judge and intra-judge reliability (Tyler, Lewis, & Welch, 2003). Overall inter-transcriber reliability was 96%.
Results

Carlos

Carlos was ~25% intelligible in conversation in Spanish. On the day of the English assessment, Carlos spoke in English only when asked to repeat the imitated sentences task, choosing to respond in Spanish at all other times. Carlos completed both the word and sentence imitation tasks for his speech assessment. Carlos spoke more clearly in single word tasks; in multisyllabic words and running speech, intelligibility decreased dramatically, primarily because Carlos produced few true consonants in longer utterances. Carlos’ intelligibility increased only slightly when he attempted to self-correct during communication breakdowns.

In the analysis of Carlos’ speech, we considered English or Spanish sounds in his phonetic inventory if they were produced in three different types of utterances across the single word and imitated sentences tasks. Carlos produced all Spanish vowel and consonant phonemes except for the alveolar trill /r/. Carlos also produced the Spanish consonant allophones [β, ǻ, ɣ]. In Spanish, Carlos produced the following English phonemes: /s/ four times, /θ/ twice and /ʃ/ and /s/ once. Carlos had a complete English vowel inventory. The English consonants that Carlos did not yet produce were /v, θ, й/; none of these sounds are phonemic in Spanish. Carlos produced the following Spanish sounds in his English: the voiced velar fricative [v] once and the voiceless velar fricative /β/ four times, /θ/ nine times. He produced a variety of English words in which the English sounds were phonemic in Spanish. Carlos produced the following Spanish sounds in his English: the voiced velar fricative [v] once and the voiceless velar fricative /θ/ nine times. He produced a variety of English words in which the English sounds were phonemic in Spanish.

Table I compares Carlos’ pre- and post-treatment accuracy and error pattern frequency in words and sentences. In the analysis of Carlos’ speech, we considered English or Spanish sounds in his phonetic inventory if they were produced in three different types of utterances across the single word and imitated sentences tasks.

*Accuracy measures*

<table>
<thead>
<tr>
<th></th>
<th>Spanish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Words</td>
<td>Sentences</td>
</tr>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>Percentage Consonants Correct</td>
<td>73.2</td>
<td>86.4</td>
</tr>
<tr>
<td>Percentage Vowels Correct</td>
<td>95.9</td>
<td>97.0</td>
</tr>
<tr>
<td>Index of Phonetic Complexity</td>
<td>1.91</td>
<td>2.32</td>
</tr>
<tr>
<td>Percentage Initial Consonants Correct</td>
<td>86.6</td>
<td>90.0</td>
</tr>
<tr>
<td>Percentage Medial Consonants Correct</td>
<td>83.3</td>
<td>85.4</td>
</tr>
<tr>
<td>Percentage Final Consonants Correct</td>
<td>44.4</td>
<td>84.4</td>
</tr>
<tr>
<td>Error rates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/s/</td>
<td>47.3</td>
<td>7.4</td>
</tr>
<tr>
<td>Stopping</td>
<td>7.6</td>
<td>1.9</td>
</tr>
<tr>
<td>Cluster Reduction</td>
<td>53.3</td>
<td>21.7</td>
</tr>
<tr>
<td>Final Consonant Deletion</td>
<td>52.8</td>
<td>9.4</td>
</tr>
</tbody>
</table>

Table I. Carlos’ pre- and post-treatment accuracy and error pattern frequency in words and sentences.

After analysis of Carlos’ pre-treatment speech characteristics, two syllable-level treatment targets were selected for treatment in Spanish and English. TRG1 was /s/+ consonant cluster (abutting consonants) production in initial and medial word position (/s/+ clusters only occur in medial position in Spanish). TRG2 was medial and final consonant and consonant cluster production, not including /s/ clusters. Word and sentence probe stimuli were developed for Carlos from age-appropriate words demonstrating the targeted errors. If a selected word had been used in treatment in the previous two sessions, a different randomly selected word from the list was probed. Prior to probe presentations, Carlos was told one time that the words contained his new sounds and that he should repeat the words and sentences the teacher said. Baseline probes were collected prior to treatment onset in the clinical setting; treatment probes were collected at the beginning of each session. Type of probe data was randomly selected each day; each type was not collected during every session.

Carlos’ baseline data were collected in both language environments as appropriate. Treatment was conducted in Spanish during sessions 5, 6, 7, 9, 10, 11, 13, 14, 15, 17, 18, and 19. Treatment was conducted in English during sessions 8, 12, and 16. Cross-linguistic generalization activities occurred at the end of every treatment session. Treatment for TRG1, /s/ clusters, began in session five after a stable baseline was achieved. After baseline, TRG1 was treated for 15 sessions for 300 minutes total. Probe data for TRG1 are shown in Figure 1 for single words and in Figure 2 for imitated sentences. As Figure 1 shows, Carlos did not produce /s/+ clusters in single words or sentences in either Spanish or English for Carlos’ baseline data were collected in both language environments as appropriate. Treatment was conducted in Spanish during sessions 5, 6, 7, 9, 10, 11, 13, 14, 15, 17, 18, and 19. Treatment was conducted in English during sessions 8, 12, and 16. Cross-linguistic generalization activities occurred at the end of every treatment session. Treatment for TRG1, /s/ clusters, began in session five after a stable baseline was achieved. After baseline, TRG1 was treated for 15 sessions for 300 minutes total. Probe data for TRG1 are shown in Figure 1 for single words and in Figure 2 for imitated sentences. As Figure 1 shows, Carlos did not produce /s/+ clusters in single words or sentences in either Spanish or English for
the first six sessions. After treatment was initiated, Spanish probe word accuracy increased to 50% in the 7th session, briefly decreasing to 30% accuracy in session 8, increasing to 70% in session 9, and stabilizing at 90% in words from sessions 16–19. As shown in Figure 2, Carlos’ production of /s/ clusters also improved in Spanish sentences, increasing in accuracy to 10% in the 8th session, 20% in session 11, 60% in session 14, 70% in session 17, and ending at 50% in session 19.

In English, treatment effects were not as rapid; however, /s/ cluster productions increased in accuracy in both words and sentences. In both environments, /s/ clusters were produced with 0% accuracy throughout the baseline period (see Figures 1 and 2). During treatment, word accuracy for TRG1 increased first in session 8 to 20%, fell again to 0% in session 9, rose to 30% in session 13, 50% in session 16, and returned to 30% in session 19. In English sentences, TRG1 accuracy remained at 0% until session 17 when it was 30%; TRG1 accuracy was 20% in session 19.

Treatment for TRG2, medial and final consonant and consonant cluster production (not including /s/ clusters) began in session 8 after a stable baseline was achieved and TRG1 was produced with 25% or greater accuracy for two sessions. TRG2 was treated for 11 sessions for 220 minutes total. Probe data for Carlos’ TRG2 are shown in Figure 3 for single words and Figure 4 for imitated sentences. As Figure 3 shows, Carlos produced medial and final consonants in Spanish with ~30% accuracy prior to treatment; this accuracy was lower in sessions 3 and 4, and decreased to the lowest level of 10% in session 6. After treatment for TRG2 was initiated in session 8, TRG2 increased to 58% in Spanish single words in session 9, hovering between 55–60% for the remaining sessions. In Spanish sentences, TRG2 was produced with 20% accuracy during the baseline phase;
After treatment was initiated, TRG2 probe data accuracy increased to 35% in session 9, decreased to 0% in session 12 and 10% in session 15, then increased to 40% and 50% in sessions 18 and 19, respectively. In English words, probe data accuracy for TRG2 was similar to Spanish, ~30%. After intervention was initiated, TRG2 single word accuracy increased to 50% in words in session 12, then increased to ~60% accuracy in sessions 14, 17, and 19. In English sentences, TRG2 accuracy for sentences was 0% during the baseline phase. TRG2 accuracy increased to 20% in session 9, returned to 0% in session 12, increased to 10% in session 15, and then increased to 40% in session 19. The effect size for baseline and treatment phases for single words was calculated using Cohen’s d. The difference in accuracy for Carlos during treatment in both Spanish and English was large for d. The accuracy increased dramatically in most situations, particularly for final consonants excluding /s/-clusters. We examined the effects of these on /s/- production and related error patterns. As shown in Table I, error frequency for the phoneme /s/ decreased in Spanish words by 40%, in Spanish sentences by 31%, and in English words by 46%. Errors on /s/ increased by 15% in English sentences. It is possible that the increase in errors in English sentences reflected a U-shaped change as suggested by Dynamic Systems Theory (DST, Gershkoff-Stowe & Thelen, 2004). In the pre-assessment, Carlos shortened many target English sentences and, by the end of therapy, Carlos was attempting the more complex English words. The frequency of cluster reduction and final consonant deletion decreased in both languages in Carlos’ sentences and words. The smallest decrease was 7% for cluster reduction in English sentences; the largest was a 42% decrease in final consonant deletion in Spanish words.

Rubén

During the assessment, Rubén spoke few sentences unless prompted; in sentences and single words Rubén was ~10% intelligible in Spanish. During English assessment, Rubén only spoke in English during the single word task; most words were attempted only after the examiner provided a model. Rubén attempted but did not successfully complete the imitated sentences task in Spanish or English; both were discontinued after multiple attempts with fewer than three or four words. Because we only have single word data and, thus, less data for Rubén, we considered a sound to be included in his phonetic inventory if he produced it two or more times in different types of utterances in the single word task. Rubén produced all Spanish vowel and all consonant phonemes except the following: the voiceless alveolar fricative /s/, the voiceless velar fricative /x/, the alveolar tap /l/, and the alveolar trill /r/. Rubén produced the Spanish consonant allophones [β], but not the [ð] or the [ʃ]. In his Spanish, Rubén produced the following English phonemes: /œ/ and /a/ once, /s/ twice, and /ʃ/ once. Rubén had a complete English vowel inventory except for the rhotic vowel /œ/. The English consonants that Rubén did not produce were /θ, z, s, j, h, ʃ, r/; of these sounds, only /s/ is phonemic in Spanish. Rubén produced the following Spanish sounds in English: the bilabial fricative [β]

Table II. Comparison of pre- and post-treatment assessment results: Rubén.

<table>
<thead>
<tr>
<th>Accuracy measures</th>
<th>Spanish</th>
<th>Post</th>
<th>English</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage Consonants Correct</td>
<td>40.2</td>
<td>47.8</td>
<td>38.6</td>
<td>54.6</td>
</tr>
<tr>
<td>Percentage Vowels Correct</td>
<td>87.2</td>
<td>91.5</td>
<td>70.6</td>
<td>83.9</td>
</tr>
<tr>
<td>Index of Phonetic Complexity</td>
<td>1.08</td>
<td>1.3</td>
<td>1.3</td>
<td>1.8</td>
</tr>
<tr>
<td>Percentage Initial Consonants Correct</td>
<td>49.3</td>
<td>66.2</td>
<td>61.9</td>
<td>76.5</td>
</tr>
<tr>
<td>Percentage Medial Consonants Correct</td>
<td>39.2</td>
<td>42.2</td>
<td>46.8</td>
<td>45.8</td>
</tr>
<tr>
<td>Percentage Final Consonants Correct</td>
<td>11.1</td>
<td>22.6</td>
<td>22.1</td>
<td>52.3</td>
</tr>
<tr>
<td>Error rates</td>
<td>/s/</td>
<td>100.0</td>
<td>88.5</td>
<td>100.0</td>
</tr>
<tr>
<td>Stopping</td>
<td>33.3</td>
<td>21.1</td>
<td>13.9</td>
<td>12.1</td>
</tr>
<tr>
<td>Final Consonant Deletion</td>
<td>77.8</td>
<td>71.0</td>
<td>48.8</td>
<td>32.6</td>
</tr>
<tr>
<td>Initial Cluster Reduction</td>
<td>85.7</td>
<td>84.2</td>
<td>76.7</td>
<td>69.0</td>
</tr>
<tr>
<td>Medial Cluster Reduction</td>
<td>66.7</td>
<td>56.9</td>
<td>45.5</td>
<td>38.9</td>
</tr>
</tbody>
</table>
once and the low central vowel /a/ 14 times. He produced primarily CV syllable shapes in both Spanish and English, although a few initial and medial clusters and final consonants were produced in both languages. Accuracy rates for consonants and vowels as well as frequency of segmental and syllabic error were analysed. Table II compares Rubén’s pre- and post-treatment accuracy and the highest frequency errors in single words affecting both languages. In addition, some language-specific errors were observed with high frequency. In Spanish, these were trill (100%) and tap (98%) substitution; gliding (33%) and derhoticization (100%) were frequent in English.

Rubén exhibited characteristics of moderate CAS in assessment and treatment: inconsistent production of consonants and vowels in repeated production of words, difficulty with motor planning, including metathesis of syllables, groping, difficulty with changing motor patterns, increased error rates in longer utterances, and unusual and frequent vowel errors. Rubén’s speech was also characterized by unusually slow rate and lengthened or abrupt transitions within syllables.

After pre-treatment speech characteristics were analysed, one segmental and one syllable-level error were selected for treatment in both languages. TRG1 was /s/. Rubén did not produce /s/ at the onset of treatment and his initial treatment focused on producing an approximation of /s/ in functional one- and two-syllable words in both languages. While the initially-planned probe sentences were collected, because of Rubén’s difficulty in /s/ production these probe words and sentences were at 0%. TRG2 was word-initial and medial stop + liquid consonant cluster production. Probe data at the single word level were collected at the beginning of each session.

Rubén’s baseline probe data were collected in both language environments. Originally probe data (baseline and treatment) were intended to be in randomly selected single words. However, since Rubén produced both TRG1 and TRG2 with 0% accuracy in probe words and in all speech prior to therapy, probe data used for comparison were altered so treatment effects could be monitored. Thus, baseline and probe data were production of TRG1 and TRG2 in one- and two-syllable single words. Phase 1 for TRG1 was percentage accuracy for single words with /s/ produced in direct imitation; Phase 2 for TRG1 was percentage accuracy for single words with /s/ in delayed imitation.

Rubén’s treatment was conducted in Spanish during sessions 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 18, 19, 22, 23, and 24. Treatment was conducted in English during sessions 8, 11, 13, 16, and 20. Rubén required intensive cueing and feedback for target productions. Rubén infrequently produced /s/ without distortions; /s/ was counted as correctly produced when it was an alveolar fricative, even if slightly distorted. Rubén exhibited motor planning difficulties, requiring visual, tactile, and occasionally auditory cues for articulatory placement. He often produced the target sounds in the wrong word position. Practice rates were delayed in therapy as needed. These treatment strategies are in line with dynamic temporal and tactile cueing, shown successful in monolingual children with CAS (e.g., Edeal & Gildersleeve-Neumann, 2011; Maas & Farinella, 2012; Strand et al., 2006). Phonological strategies were also included in therapy. Metaphonological activities to encourage cross-linguistic generalization occurred at the end of every treatment session.

Treatment for TRG1, initial and final /s/ singletons, began in session four after achieving a stable baseline. TRG1 was treated for 20 sessions for 400 minutes total. Figure 5 shows probe data for single words. There were two phases of probe data for TRG1 treatment: direct imitation (phase 1) and delayed imitation (phase 2). As Figure 5 shows, Rubén produced /s/ with 0% accuracy during baseline and through session 6 in both languages. Starting in session 7, /s/ production increased in accuracy in direct imitation in Spanish and English, staying above 50% accuracy starting in session 14. Production accuracy increased more rapidly in Spanish than in English. Starting in session 19, probe data were collected for target utterances produced in delayed imitation. After an initial slight decrease in accuracy in Spanish, Rubén produced target words with at least 75% accuracy in both languages.

Treatment for TRG2, stop + liquid cluster production, began in session 19, at the start of Phase 2 for TRG1. TRG2 was treated for seven sessions for a total of 124 minutes of treatment. Probe data for Rubén’s TRG2 are shown in Figure 6 for single words. Rubén produced stop + liquid clusters with 0% accuracy in Spanish and English during baseline; target words were produced with between 10–20% accuracy in treatment sessions 2–4 in both languages. In sessions 5–7, TRG2 words were produced with 60% accuracy or greater. Effect size for baseline and treatment phases for single words.

Figure 5. Rubén’s probe data measurements for TRG1: Accuracy rates for /s/ production in one- and two-syllable words.
for Rubén’s was calculated using Cohen’s $d$. The difference in accuracy for Rubén’s during treatment in both Spanish and English was large for TRG1 ($d = 2.355$), TRG2 ($d = 3.173$) and overall ($d = 2.644$).

A comparison of Rubén’s pre- and post-treatment speech samples is shown in Table II. PCC increased by 8% in Spanish and by 16% in English. PVC increased by 4% in Spanish and 13% in English. In words and sentences in both languages, phonetic complexity (frequency of later-developing sounds and complex word shapes) increased. The percentage of initial consonants produced correctly increased by 17% in Spanish and 14% in English. The percentage of final consonants produced correctly increased by 12% in Spanish and 30% in English. Medial consonants did not change in English and increased in accuracy in Spanish by 3%.

The frequency of error in the phoneme /s/ and error patterns affected by the treatment targets was compared. Error frequency for /s/ decreased by 11.5% in Spanish, from 100% to 88.5%, and 16% in English, from 100% to 84%. Stopping decreased by 11% in Spanish and almost 2% in English. Final consonant deletion decreased from 78% to 71% in Spanish and from 49% to 33% in English. Also, initial and medial cluster reduction decreased in post-assessment (although not treated, many of these clusters also include /s/). Initial cluster reduction did not change much in Spanish, remaining at ~85%, but decreased by 7% in English, to 39%. Medial cluster reduction decreased from 67% to 57% in Spanish and from 46% to 39% in English.

Discussion

This study explored the effectiveness of bilingual therapy in two bilingual Spanish–English boys with SSDs. We hypothesized that selecting targets that affected both languages at relatively high rates and providing intervention in both languages would result in cross-linguistic generalization, measured by significant increases in phoneme accuracy and phonetic complexity and significant decreases in the percentage-of-occurrence of error patterns in both languages. The hypothesis supports a dynamic systems perspective to speech development. The multiple-baseline across behaviours design provided a controlled experimental condition for each child’s response to intervention.

Our study shows that bilingual therapy for SSDs targeting phonological representation, articulation, and metaphonological skills can be effective in improving the child’s overall sound system. Results indicated that the children’s speech improved in both languages, with a medium-to-large effect size depending on treatment target observed in both children. Although the accuracy of treated targets improved in all cases, increases in overall accuracy were also noted, particularly overall consonant and vowel accuracy and utterance-level complexity. Gains were similar in magnitude in both languages, even with the Spanish language emphasis in treatment. We recognize that our research design does not allow for efficiency measures, leaving us unable to determine the strength of the gains with an English-only or predominantly-English treatment model. Nevertheless, the reports from previous therapy (discussed below), in which the children made limited gains in their primarily English treatment, lead us to believe that focusing on the children’s first language and their stronger language resulted in greater overall treatment gains than an emphasis on their weaker language would have provided.

We also found that our treated targets resulted in improvement in non-treated phonological error patterns, such as final consonant deletion, which decreased in both children, particularly in English. While one might interpret these changes as evidence of maturity, external factors suggest this is not the case. Both Carlos and Rubén had participated in treatment for 6 months prior to the study. Their previous therapy was once per week in small groups. No control of number of productions per session was calculated. Parent and school therapist progress reports suggest little to no improvement in speech production goals during this time period. Also, TRG2 accuracy for both children did not change in error frequency during the control periods. We propose that the many positive effects on non-treated behaviours (e.g., decreases in non-target patterns, increases in PVC and PCC in English and Spanish) in our treatment in Spanish and English is in keeping with a dynamic systems perspective and results from cross-linguistic generalization of treatment, treatment frequency and session length, and the number of target productions per session.

Five minutes of each treatment session were spent demonstrating to Carlos and Rubén how the speech characteristic practiced that day in one language
also applied to properties of the other language. The importance of meta-phonological tasks in treatment for bilingual SSDs has not been explored previously, and it is difficult to know if the improvement of targets in both languages resulted from the meta-phonological generalization, the dual-language treatment, or both. Further studies should explore these possibilities.

One unexpected finding was the increase in vowel accuracy in both children. This change was more dramatic in English. While we did not specifically treat vowels nor did we assume that vowel error was an aspect of either child’s speech sound disorder, it appears that the clear treatment models provided in English may have helped both boys build more accurate perceptual and production representations for English vowels (Rvachew & Nowak, 2004).

Our treatment emphasized Spanish, with approximately three of four sessions conducted in Spanish. The decision to primarily focus on Spanish was motivated by the sequential bilingual situation of both children—they were in Spanish home settings, spent time primarily with family and friends who speak Spanish, and demonstrated stronger receptive and expressive Spanish skills. As noted in the results section, Rubén had both a segmental goal, the phoneme /s/, and a syllable-level goal, the production of stop + liquid clusters. Both goals for Carlos addressed phonotactic complexity, particularly in longer utterances. Goals were chosen for multiple reasons: functional load in both languages, frequency in both Spanish and English, and developmental appropriateness. It is important to consider the importance of the targets you select in a bilingual child as they apply across languages.

Results showed that targeting associations between languages results in cross-linguistic generalization and suggest the interconnectedness of seemingly autonomous languages, an inter-connectedness specified by DST (Thelen & Smith, 1995). In these two children, such generalization resulted in their ability to use skills in one language for more accurate productions in the other (Paradis, 2001). Both children showed improvement in both languages, not only on target structures but also on their phonological systems as a whole. For example, both overall consonant and vowel accuracy increased from pre- to post-treatment, indicating system-wide changes that went beyond improvement on the treatment targets themselves (although accuracy on those targets improved as well). Such target-specific and system-wide change shows at least some evidence for DST in that phonological properties which are closely associated result in greater cross-linguistic generalization across languages. This cross-linguistic generalization has been noted in studies examining the phonological skills of Spanish–English (e.g., Fabiano-Smith & Goldstein, 2010; Gildersleeve-Neumann et al., 2009) and Russian–English (Gildersleeve-Neumann & Wright, 2010) bilingual children. Albeit limited at this point, there is converging evidence from treatment and non-treatment studies supporting cross-linguistic generalization, as would be indicated by Dynamic Systems Theory.

Our study of bilingual intervention showed that treating both languages had an overall positive effect on the speech system of sequential bilingual children with SSDs. We observed fairly equal change in Spanish and English, even though the ratio of therapy was ~3:1, Spanish-to-English. Thus, treating the stronger language appeared to have a greater impact on the children’s overall system. We also found that an increase in English vowel accuracy in at least some sequential bilingual children can occur without targeted vowel therapy.

As a postscript, Rubén’s and Carlos’ changing linguistic environments highlight the bilingual child’s shifts in opportunities, necessities for both languages throughout development, and emphasize the importance of considering the whole child in treatment language choices. Both children continued in the clinic for a year after this study. Rubén’s family then moved 1 hour away and he left our clinic. However, after 6 months Rubén’s family returned, driving twice a week after school for services. Rubén’s parents believed his English-only school setting and English-only therapy were resulting in loss of speech skill in Spanish and English. Rubén’s teacher noted after his return to bilingual therapy that his English speech skills improved. Carlos also moved: his family returned to Mexico a year after his treatment ended. Although we have lost contact, we are happy that Carlos received services in Spanish, since he currently relies on his Spanish more than English skills in his educational and social communication. For many bilingual children, the home language is often more important than English in many communication settings and can have greater importance in future situations.

Our results show that bilingual treatment was effective in treating speech sound disorder in two Spanish–English children. Our treatment language ratio emphasized Spanish over English and resulted in improvement in both languages. This Spanish emphasis was based on the ages, length of exposure to English, and home and school profiles of Carlos and Rubén. The ratio will not be the same for all bilingual children or all sequential bilingual children; instead, language histories for each client will need to be considered in determining the ideal ratio.

Our study is limited by its scope; there are no other treatment data for bilingual Spanish–English children with SSD. Clearly, extensive treatment efficacy research needs to be completed to understand the best language environments for treatment of bilingual children. Bilingual children differ in multiple ways and the language needs in treatment will vary. The understanding of intervention for bilingual children is in its infancy. A variety of studies are needed to determine how treatment needs differ for
simultaneous and sequential bilinguals, for different language typology combinations, and for children who have concomitant disorders, among other factors. Research should include more single-subject design studies to explore change in individual children and randomized control studies to explore group treatment effects. To better serve bilingual children, we need to understand the effectiveness as well as efficiency of English-only, Spanish-only, and bilingual therapy for bilingual children.

**Note**

1. IPC of each word is derived by calculating the point value for each of eight complexity factors (consonant place, consonant manner, vowel type, word length, word shape, cluster production, type of cluster, and intra-word consonant sequencing). IPC has been validated in English (Howell, Au-Yeung, Yaruss, & Eldridge, 2006) and Spanish (Howell & Au-Yeung, 2007).

**Acknowledgments**

We are grateful to Claudia Meyer and Lisa Arakelian for their contributions to treatment and to the families who took part in this study.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

**References**


Cross-linguistic generalization in speech treatment


