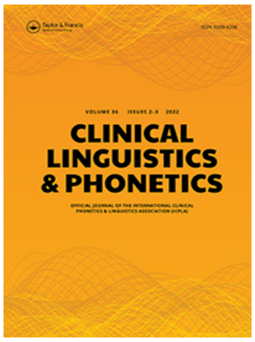

Clinical Linguistics & Phonetics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/iclp20

Phonological complexity in intervention for Spanish-speaking children with speech sound disorder

Philip Combiths, S. Pruitt-Lord, A. Escobedo & J. A. Barlow

To cite this article: Philip Combiths, S. Pruitt-Lord, A. Escobedo & J. A. Barlow (2022) Phonological complexity in intervention for Spanish-speaking children with speech sound disorder, Clinical Linguistics & Phonetics, 36:2-3, 219-240, DOI: 10.1080/02699206.2021.1936186

To link to this article: https://doi.org/10.1080/02699206.2021.1936186

Published online: 10 Jun 2021.

CLINICAL LINGUISTICS & PHONETICS
2022, VOL. 36, NOS. 2–3, 219–240
https://doi.org/10.1080/02699206.2021.1936186

Phonological complexity in intervention for Spanish-speaking children with speech sound disorder

Philip Combiths[a,b], S. Pruitt-Lord[a], A. Escobedo[a,b], and J. A. Barlow[a]

[a]School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, California, USA; [b]Department of Cognitive Science, University of California, San Diego, California, USA

ABSTRACT

The efficiency of intervention for children with speech sound disorder may be influenced by linguistic complexity of the phonological intervention target. Complex targets, particularly, later-acquired, less-known consonants and consonant clusters, have been linked to greater post-intervention generalization to untargeted phonological structures. Yet there is little direct evidence to support target selection based on linguistic complexity for Spanish-speaking children with speech sound disorder. This intervention study utilizes an experimental single-case design to examine the efficacy of intervention in Spanish using different complex targets (i.e. /gɾ/, /bɾ/, and /l/). For each of the four Spanish-speaking children with speech sound disorder, sounds at 0% accuracy during baseline were monitored across the baseline period, during and post-intervention, and at one- and two-month follow-up visits. Over the course of intervention, only one participant achieved mastery of the targeted structure in practiced words. However, all participants demonstrated some amount of broad phonological generalization to untargeted consonants or clusters. Variable learning trajectories and broad phonological generalization are discussed as they relate to participant characteristics and linguistic complexity.

ARTICLE HISTORY
Received 28 September 2020
Revised 23 May 2021
Accepted 24 May 2021

KEYWORDS
Phonological disorder; speech sound disorder; Spanish; bilingual; intervention; treatment

Introduction

The heterogeneous population of children with communication impairments requires access to equitable, evidence-based, and efficient interventions so that they are prepared for better language outcomes and academic success (Almost & Rosenbaum, 1998; Law et al., 2004). Spanish is the second most widely spoken native language in the world, surpassed only by Mandarin (Fernández Vítores, 2017), yet there is a paucity of research available to support intervention decisions for Spanish-speaking children with speech sound disorder (SSD), one of the most prevalent categories of communication disorders in young children (Law et al., 2000). In short, our evidence base for speech interventions in non-English languages is decades behind what has been accumulated for English. In response to this disparity, we must conduct translational research sampled in a way that is more representative of the diversity of children with SSD to provide evidence-based guidelines for speech intervention in Spanish and other non-English languages.

CONTACT Philip Combiths pcombiths@sdsu.edu School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, CA 92182

Converging evidence, primarily from populations of monolingual English speakers with various communication disorders, has indicated an advantage in the efficiency of interventions which target more difficult, challenging, or complex components of language (e.g. Gierut, 1999; Thompson et al., 2003; Van Horne et al., 2017). For children with SSD, this type of intervention target selection has been referred to as a complexity or complexity-based approach. Evidence supporting a complexity-based approach (e.g. Elbert et al., 1984; Gierut, 1990, 1998a, 1999; Gierut & Morrisette, 2012; Gierut et al., 1996; Pagliarin et al., 2009; cf. Rvachew & Bernhardt, 2010) suggests that the optimal intervention target is a phonological feature, contrast, or structure that is typologically less common across languages (e.g. Ladefoged & Maddieson, 1996), later developing in children (e.g. McLeod & Crowe, 2018), and less known by the individual (e.g. Gierut et al., 1987). These targets are thus linguistically complex as well as relatively complex for a given child (Gierut, 2007). This approach is motivated by implicational relationships that explain cross-linguistic tendencies for phonological systems with given complex structures to obligatorily also include or *imply* certain simpler structures (e.g. Ladefoged & Maddieson, 1996). This, in conjunction with principles of language learnability (Pinker, 1984), provides a framework for the observation that intervention targeting such complex structures tends to result in broad, across-class generalization to untargeted sounds (Gierut, 2007). In other words, children trained with complex exemplars may demonstrate improvement beyond the targeted structure, specifically in related structures of similar complexity and cascading to other less complex structures.

Paradigms for implementing a complexity-based speech intervention (e.g. Baker & Williams, 2010; Storkel, 2018) commonly target relatively late-acquired singleton consonants (e.g. /l/, /s/ in English) or consonant clusters (e.g. /fl/, /stɹ/in English). Relative complexity among clusters has been associated with sonority, a phonological construct correlated with acoustic intensity (per the sonority sequencing principle; Clements, 1990; Parker, 2002), such that clusters with a smaller sonority distance between consonants are less frequently occurring and more complex. As highlighted, most of the evidence base for complexity-based target selection has been limited to monolingual English-speaking children. However, linguistic complexity is a language-general phenomenon, with some language-specific idiosyncrasies. Thus, the general principles that have guided the existing research with English-speaking children should apply to speakers of other languages (Watts & Rose, 2020). This has been generally attested in two case studies targeting complex clusters in Spanish (Anderson, 2002; Barlow, 2005) and investigations with Portuguese-speaking children (Barberena et al., 2015; Ceron et al., 2013; Mota et al., 2007; Pagliarin et al., 2009; Pereira & Mota, 2002). However, the language-specific characteristics of Spanish phonology mean that ideal complex targets for achieving broad generalization in Spanish will not be identical to those of English (e.g. Cataño et al., 2009).

Spanish phonology diverges from English phonology in many areas relevant to complexity-based target selection. For instance, even for cross-linguistically similar phonemes, such as /l/ or /s/ in English and Spanish, age of acquisition can differ substantially (McLeod & Crowe, 2018), and this may relate to the relative complexity of those segments within each language. Furthermore, English and Spanish differ in their repertoires of permitted consonant clusters. English permits many onset clusters with up to three elements (e.g. /spl/), whereas Spanish has a more limited distribution of two-element onset clusters (e.g. /pl/). Importantly, more than half of the potential consonant clusters in Spanish are consonant

+glide clusters (e.g. /bw/, /fj/) (Real Academia Española, 2011), and the status of these sequences as "true" branching onset clusters in Spanish is debated (Barlow, 2005; Harris, 1983; Senturia, 1998). Highly sonorous glides (e.g. /w/, /j/) have vowel-like properties and may position in the syllable nucleus in some instances or some languages (Blevins, 1995). In Spanish, glides that are the second element in a syllable-initial cluster (e.g. /pwente/*puente*, "bridge") have been posited either in the nucleus as part of a diphthong (e.g. Harris, 1983) or in the onset as part of a consonant cluster (e.g. Senturia, 1998). It may also be the case that the position of glide segments in syllable structure is dynamic and variable within and across children during Spanish acquisition (Barlow, 2005). Given conflicting evidence, the syllable constituency of glide segments in a cluster is not clear, which has implications for the relative complexity of consonant+glide sequences (for additional discussion of Spanish phonological complexity, see Barlow, 2003; Cataño et al., 2009).

Despite cross-linguistic differences in the parameters of complexity that would impact complexity-based target selection, there is little evidence with which to compare phonological generalization patterns across potentially complex intervention targets in Spanish or any language other than English. To support intervention for Spanish-speaking children with SSD, we must better examine the impact of phonological complexity on intervention provided in Spanish.

Current study

In this initial study, we employ a single-case experimental design to examine the effect of speech intervention targeting relatively complex phonological structures in Spanish, including complex clusters (e.g. /ɡɾ/) and complex singletons (e.g. /l/) in four Spanish-speaking bilingual children with SSD. Given the paucity of research regarding the phonological characteristics of intervention targets for SSD in Spanish, we address the following questions:

- Is phonological learning stimulated by intervention targeting relatively complex phonological structures?
- What are the observable patterns of phonological generalization following intervention targeting complex singletons and clusters in Spanish?
- Clinically, we contribute to the limited body of intervention efficacy research for Spanish-speaking children with SSD and identify individual characteristics and features of the intervention targets that may impact a child's response to this speech intervention.

Method

This study is part of an NIH-funded clinical trial (clinicaltrials.gov, No. NCT03977701) and was approved by the Institutional Review Board and Human Subjects Protection Program at San Diego State University. The study procedures were explained to parents and participating children in person. Parents then provided written informed consent, and the children assented to their participation.

Participants

Four Spanish-English bilingual children (hereafter referenced with pseudonyms), aged 4:1–5:11, with a phonologically based SSD, also referred to as functional phonological disorder (Gierut, 1998b), are included in this study as participants from a larger, ongoing intervention study. Participants were recruited by circulating digital and paper fliers at a university speech-language clinic and to local speech-language pathologists. The participants are first-language speakers of Spanish living in Spanish-dominant households. Living in the US, they are also second-language learners of English. All participants had some exposure to English through acquaintances, siblings, or preschool, and two participants (Jaime and Roberto) had a parent who was proficient in both Spanish and English. Each participant's relative exposure is quantified as a ratio of years (or portions of years) of Spanish exposure to years of English exposure since birth, as displayed in Table 1. For instance, Marta was exposed to Spanish from birth (4.1 years) and English for one year in preschool. Her ratio of Spanish: English exposure at the time of the study was thus 4.1:1 (4.1). That these children are bilingual is not a focal point of this study, nor is this study an examination of cross-linguistic intervention effects. Although some exposure to English limits our ability to compare Spanish and English as discrete linguistic entities, we gain the ability to independently examine an intervention effect in an understudied group that is more representative of young Spanish-speaking children in majority English-speaking countries.

The presence of phonological disorder was determined via a converging approach (Restrepo, 1998), including reported concern with speech development or intelligibility in Spanish by a parent and the study speech-language pathologist, in addition to the absence of 10 or more Spanish phonemes or consonant clusters from their phonetic or cluster inventories. Participants had age-appropriate language comprehension in Spanish (Zimmerman et al., 2012), normal hearing, nonverbal cognition within the normal range (Roid & Miller, 1997), performance within the normal range on an oral-motor examination, and no diagnosis of other motor, behavioural, cognitive or neurological impairment at the time of their participation in the study. Additional participant characteristics are given in Table 1, and inclusionary and exclusionary criteria are specified in Table 2.

Assessment procedures

The Evaluación de la Fonología Española (Barlow & Combiths, 2019) and individualized subsets of this probe used at baseline and mid-intervention timepoints were the primary

Table 1. Participant characteristics at initial assessment.

Name	Age	Sex	Spa:Eng Exp.*	Mat. Ed.	PLS-5 AC/EC	MLUw	Leiter-R	PCC-R	ICS
Diego	04;09.27	M	6.3	HS	84/87	2.74	11.5	57%	2.1
Jaime	05;11.15	M	3.1	HS	96/107	2.65	14	54%	2.1
Roberto	04;05.20	M	3.1	college	120/115	2.25	16	63%	2.6
Marta	04;01.10	F	4.1	HS	104/87	2.32	11	54%	2.8

Mat. Ed. = maternal education. PLS-5 = Standard scores from the Preschool Language Scales-Fifth Edition Spanish (Zimmerman et al., 2012). AC = Auditory Comprehension subtest. EC = Expressive Communication subtest. MLUw = mean length of utterance in words (Spanish). Leiter-R = Scaled scores from the Figure Ground and Form Completion subtests of the Leiter International Performance Scale–Revised (Roid & Miller, 1997). PCC-R = Percentage of Consonants Correct-Revised (Shriberg et al., 1997). ICS = Intelligibility in Context Scale (McLeod et al., 2012). ICS is displayed as total average score (maximum 5).

* Ratio of years of Spanish exposure to years of English exposure since birth.

Table 2. Inclusionary and exclusionary criteria.

Inclusionary criteria
Language profile
Mexican-US Spanish as first language
Living at home with a native Spanish-speaking caregiver
Phonological disorder
≥ 10 phonemes or clusters missing from phonetic and cluster inventories
Converging parent and SLP report of speech concern (Restrepo, 1998)
Exclusionary criteria
Receiving other speech/language services
Diagnosis of other motor, behavioural, cognitive or neurological impairment
Binaural hearing screen failure
Atypical oral-motor examination
Leiter-R nonverbal cognition standard score < 77.5
PLS-5 Spanish AC standard score < 77.5

PLS-5 Spanish AC = Preschool Language Scales-Fifth Edition Spanish (Zimmerman et al., 2012) Auditory Comprehension subtest. Leiter-R = Leiter International Performance Scale-Revised (Roid & Miller, 1997) Figure Ground and Form Completion non-verbal intelligence subtests.

sources of speech production data in this study. The EFE is a picture-based single-word elicitation probe for phonological analysis designed to sample all consonants, consonant clusters, and vowels of Spanish a minimum of three times in each permissible word position (see appendix for word list). To preserve this as a generalization measure, words in the EFE were never used during intervention. The EFE exists in A and B versions, each with a unique set of images and cues presented in different orders. Repeated administrations of the EFE with the same participant alternated between A and B versions. A subset of the EFE target words was used to create individualized probes for each of the participants' monitored sounds, sampling each a minimum of three times. This subset probe was administered at baseline and mid-intervention sessions, as described in the following section.

Dependent and descriptive variables

For each child, following analysis of their pre-intervention productions from the EFE, consonant phonemes and consonant clusters produced with 0% accuracy (henceforth 0%-at-baseline sounds) were monitored across baseline and during and after intervention. Each participant's 0%-at-baseline monitored sounds are displayed in Table 4. These monitored structures provide data for the dependent variable in this study: accuracy of 0%-at-baseline sounds (Gierut et al., 2015). Monitoring 0%-at-baseline sounds is desirable for measuring generalized phonological change because these sounds are less likely to improve on their own in a short period of time (i.e. the 6-week intervention time frame; Dinnsen & Elbert, 1984; Miccio et al., 1999; Powell, 1993; Powell et al., 1991; Sommers et al., 1967). Because this study examines generalization across the phonological system to untreated structures, it was critical to identify structures that would remain stable in the absence of intervention (i.e. <10% variability across baseline sessions; McReynolds & Kearns, 1983). A stable dependent measure was necessary to isolate an intervention effect by mitigating the confounds of time, maturation, and general variability in performance across probes.

There is also precedent for the sensitivity of composite change in accuracy of difficult sounds as an outcome measure in speech intervention research. These accuracy measures may be better differentiators of intervention effects across groups or conditions (Smit et al., 2018) than traditional global accuracy measures, such as Percentage of Consonants Correct-Revised (PCC-R; Shriberg et al., 1997). Further, much of the work examining the impact of complex target selection in treatment of phonological disorders for monolingual English-speaking children has utilized monitoring of low- or zero-accuracy sounds to establish baseline stability and operationalize broad phonological growth (e.g. Elbert & McReynolds, 1985; Gierut, 1999; Rvachew & Bernhardt, 2010).

In addition to the experimentally controlled dependent variables derived from each participant's subset of 0%-at-baseline sounds, additional measures were derived from the children's speech productions or otherwise collected across study phases. These descriptive measures included accuracy of the intervention target in practiced word (i.e. non-generalized learning), phonetic and cluster inventories, PCC-R (Shriberg et al., 1997), and parent report of intelligibility across contexts (McLeod et al., 2012). All assessment measures are listed in Table 3.

Experimental design

This study uses an experimental single-case design with multiple baselines (Byiers et al., 2012; Gierut, 2008; Kratochwill et al., 2010; McReynolds & Kearns, 1983) suited to the study population and research questions as follows. First, this design does not assume homogeneity across participants and is appropriate for intervention research for children who are highly variable in terms of speech and language use (Shriberg & Lof, 1991). Second, the design lends itself to multiple measurements and a breadth of data, which are descriptive of

Table 3. Assessment measures by study phase.

Phase	Assessment Measure
Pre	EFE
	Little PEEP
	Stimulability
	Language sample (Spanish)
	Language sample (English)
	ICS
	PLS-5 Spanish
	Hearing screening
	Oral/peripheral mechanism exam
	Developmental and language history questionnaire
	Leiter-R
Baselines 1–4	EFE (monitored subset)
Mid	EFE (monitored subset)
Post	EFE
	Little PEEP
	Language sample (Spanish)
	Language sample (English)
	ICS
1 Month Post	EFE
2 Month Post	EFE

EFE = Evaluación de la Fonología Española (Barlow & Combiths, 2019). Little PEEP = Shorter Protocol for the Evaluation of English Phonotactics (Barlow, 2012). ICS = Intelligibility in Context Scale (McLeod et al., 2012). Stimulability = Spanish stimulability task adapted from Glaspey and Stoel-Gammon (2005). Leiter-R = Leiter International Performance Scale-Revised (Roid & Miller, 1997). PLS-5 = Preschool Language Scales Fifth Edition Spanish (Zimmerman et al., 2012).

Table 4. 0%-at-baseline sound structures monitored for generalization.

Participant	Monitored sound structures
Diego	[bl, bɾ, dɾ, fl, fɾ, ɡl, ɡɾ, kl, kɾ, pl, pɾ, tɾ, lj, lw, rj, rw, mj, nj, nw, bj, fj, pj, sj, tj, r]
Jaime	[bɾ, dɾ, fɾ, ɡɾ, kɾ, pɾ, tɾ, lw, rj, rw, bw, j̞ w, r]
Marta	[bl, bɾ, dɾ, fl, fɾ, ɡl, ɡɾ, kl, kɾ, pl, pɾ, tɾ, lj, rj, mw, nj, nw, fj, j̞ w, sj, r]
Roberto	[bl, bɾ, dɾ, fl, fɾ, ɡl, ɡɾ, kl, kɾ, pɾ, tɾ, lj, rj, mw, nw, bw, dw, fw, j̞ w, kw, pw, tw, r]

each individual's response to intervention. Third, participants' baseline stability within and across conditions provides a point of reference against which the intervention effect can be observed. Finally, this design is appropriate in cases where the dependent variable (i.e. phonological accuracy) is not likely to be reversed after intervention is withdrawn (Kratochwill et al., 2010) as has been attested in prior intervention research for children with SSD (Gierut et al., 2015).

Per this design, participants were monitored across four sessions during a baseline period of no intervention, collectively demonstrating stability of participants' monitored sets of 0%-at-baseline sounds and their intervention targets in the absence of intervention to be contrasted with the observed intervention effect over the course of the intervention period and immediately post-intervention.

Intervention targets and materials

As part of their participation in the larger study, each child was randomly assigned to receive either a complex singleton or cluster target. Within these constraints, target selection was based on each child's phonological system at the pre-intervention assessment. For each child, consonants or clusters used with less than 30% accuracy were considered potential targets. This was motivated by prior research which identified 20–30% accuracy as a range of potential cut-off values, below which functional contrastive use is less likely (Combiths et al., 2019). Of these potential targets, that which was most complex, according to age of normative acquisition for singletons (McLeod & Crowe, 2018) and sonority distance for clusters (Clements, 1990), was selected as the participant's intervention target. In instances where multiple relatively complex targets were plausible, the structure with the lowest accuracy was selected. Targets were excluded if they were limited by positional constraints (i.e. /ɾ/). Potential singleton targets with multiple or complex gestures (i.e. /r/ and /ʧ/) were also excluded due to articulatory characteristics that complicate the distinction between singleton and cluster targets (e.g. Berns, 2013). Following these procedures, the cluster targets /ɡɾ/ and /bɾ/ were selected for Jaime and Diego, respectively. The target /l/ was selected for both Marta and Roberto.[1] Production accuracy of each participant's target did not vary more than 10% during baseline (McReynolds & Kearns, 1983).

Phonological targets were embedded in six words used throughout the intervention (see appendix for treatment word list). These included three real Spanish words (e.g. *grupo* "group", *grano* "grain", and *grave* "serious/bad") and three nonwords following the

[1]Both participants with singleton targets were trained with /l/. This consonant is considered less complex in Spanish relative to English (McLeod & Crowe, 2018); however, its relative complexity may be different in the unique phonological system of a bilingual child (Fabiano-Smith & Goldstein, 2010). Nevertheless, it was the least accurate, most complex singleton available for target selection in both cases.

phonotactic restrictions for permissible words in Spanish (e.g. *graki, gruka*, and *grema*). Nonwords were included as they have been shown to more readily induce generalization in children with SSD (e.g. Cummings & Barlow, 2011). Associated with these words were a set of picture cards, a story embedded with the target words, and a simple toy associated with each noun. All additional toys and materials nonspecific to the target words were consistent across participants.

Intervention protocol

Intervention was provided in one-to-one sessions for 45 minutes, three times per week, for a maximum of six weeks. All intervention was provided by one Spanish-English bilingual speech-language pathologist and conducted in Spanish following a drill-play format (Shriberg & Kwiatkowski, 1982) with an imitation phase followed by a spontaneous phase. Participants completed the intervention by attending 18 sessions or by meeting performance criterion. Criterion for completion of the imitation phase was 75% accuracy on the treatment probe across two consecutive sessions. Criterion for the spontaneous phase was 90% accuracy on the treatment probe across three consecutive sessions (Gierut, 2015).

At the start of intervention, children were oriented to their target with visual, verbal, tactile, and articulatory cues to achieve stimulability of the target or a close approximation (Bauman-Waengler, 2008; Secord, 2007). In each session, the clinician attempted to maximize the child's target production attempts (M = 226, SD = 63). During the imitation phase, productions were elicited in imitation with 1:1 clinician feedback. During the spontaneous phase, elicitations did not include a verbal model, and spontaneous production was facilitated. Feedback was intermittent during this phase to allow for self-monitoring and self-correction (Ertmer & Ertmer, 1998; Shriberg & Kwiatkowski, 1990).

Transcription and analysis

At the beginning of each intervention session, the target accuracy probe was administered, eliciting each of the child's target words three times via images displayed on a tablet. No feedback was provided during the target accuracy probe. Productions in the target accuracy probe were scored by the clinician as accurate or inaccurate, based only on production of the target. Thus, age-appropriate production of the targeted singleton or cluster was scored as accurate and any other production of the target (including complete omission) was scored as inaccurate, independent of the child's production of the rest of the word.

All participant productions from the EFE at each timepoint were transcribed online with narrow phonetic notation by the administering clinician or a Spanish-speaking research assistant trained in Spanish transcription and recorded onto a Roland Edirol R-09 digital recorder at a sampling rate of 44,000 Hz for later transcription by a different research assistant blind to the original transcriptions. Point-to-point interrater reliability for 20% of each session was 90.4% (e.g. Shriberg & Lof, 1991). Transcriptions were entered in Phon (version 2.2; Rose & Hedlund, 2017), after which production targets were generated and aligned to each child's actual productions. For the purposes of analysis, all onset consonant sequences, including consonant+glide sequences, were categorized as consonant clusters. Accuracy for 0%-at-baseline consonants and clusters and PCC-R (Shriberg et al., 1997) were generated within Phon. Phonetic and cluster inventories based on a two-time occurrence in productions from the EFE were generated with the AutoPATT plugin (Combiths et al., in press).

Results

In this section, we provide the results of the study, beginning with a qualitative description of intervention components and each participant's intervention progress in terms of accuracy with the targeted structure in practiced words. Then, we describe the dependent variable, which captures broad phonological generalization to 0%-at-baselines sounds, and provide an estimate of the intervention effect size. Finally, we present additional descriptive measures of each participant's phonological system and intelligibility.

Intervention and target accuracy

A retrospective qualitative analysis of intervention sessions according to the phonological intervention taxonomy proposed by Baker et al. (2018) confirmed that teaching moments were primarily articulatory-phonetic and phonological in nature, and most sessions included both. Spoken cues and models were provided in every session, and visual or gestural models were used in some sessions for all participants. Tactile cues were used infrequently in some sessions for Diego and Roberto only, as follows: for Diego, finger tapping or moving two fingers together was used to emphasize the presence of two components in his cluster target, /bɾ/; for Roberto, the area around the nose was touched to remind him to produce his target /l/ without nasal airflow.

Children's production attempts were imitated, elicited without a model, or spontaneous, with elicited attempts introduced after the midpoint of intervention. Every session included production attempts in real words and nonwords, and few production attempts were made in isolation. Sessions after the midpoint of intervention included more productions in sentences and conversation. Feedback given by the clinician following production attempts primarily offered knowledge of results (i.e. correct/incorrect; e.g. "¡así es el nuevo sonido!" [that's your new sound!]) or knowledge of performance (e.g. "la próxima vez con los dos sonidos juntitos" [next time with both sounds together]). Knowledge of results and performance was given in nearly every session. Clinician recasts demonstrating an accurate production were also frequent, occurring in about half of sessions. Only few instances of explicit self-reflective feedback were recorded. Finally, session activities combined drill and play, with interactive reading included after the midpoint of treatment.

Each participant's performance on the target accuracy probe for each intervention session is displayed in Figure 1. These data reflect the participants' production accuracy of their treatment target, in three productions of each of their treatment words (18 total) collected at the beginning of each treatment session. One participant, Marta (target /l/), met performance criteria for completion at session 15. The remaining participants completed 18 intervention sessions and demonstrated considerable variability in their learning of the targeted structure. Jaime (target /gɾ/) exhibited large fluctuations, ending at 61% accuracy with his target. Roberto (target /l/) also fluctuated, falling near or below 10% target accuracy until session 16 where accuracy rose sharply, ending at 89%. Diego (target /bɾ/) did not

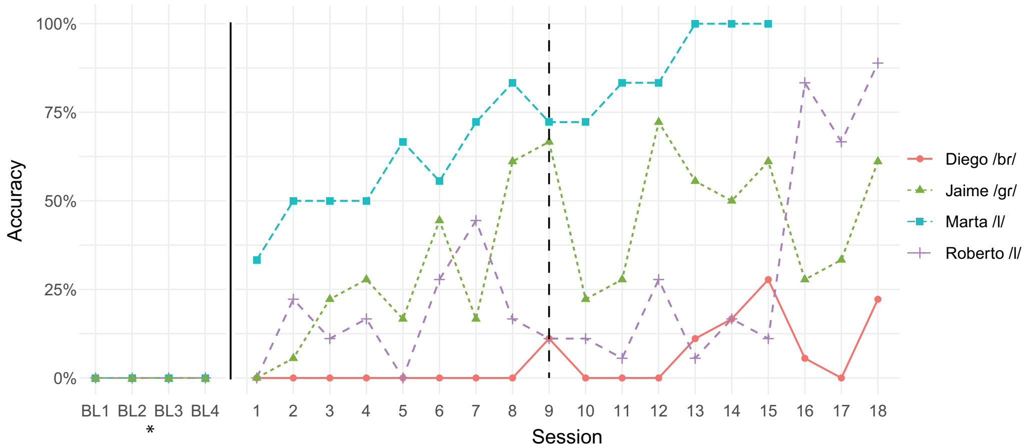

Figure 1. Intervention target accuracy. Solid line separates baseline sessions (pre-treatment) and treatment sessions. Dashed line indicates the treatment midpoint/phase shift. BL = Baseline session.* Treatment target accuracy at baseline sessions is derived from productions of each participant's target structure in onset position from baseline probes (i.e. not treatment words).

produce any accurate instances of his intervention target until session 9 and completed the intervention at only 22% accuracy.

Generalization to monitored sounds

Considering each child's monitored set of 0%-at-baseline singleton consonants and clusters only, we can conservatively examine patterns of change most attributable to an intervention effect. Composite accuracy for 0%-at-baseline consonants and clusters is displayed in Figure 2. Note that all monitored sounds remained stable at 0% accuracy across the baseline sessions. Jaime demonstrated the largest amount of growth in 0%-at-baseline structures, although fluctuating considerably. Diego and Marta showed smaller but more consistent patterns of broad growth. Roberto showed the most delayed pattern of generalization, with no change in 0%-at-baseline structures at the midpoint of intervention, and some growth evident at the post-intervention assessment.

Each 0%-at-baseline singleton or cluster for which accuracy changed over the course of the intervention is displayed in Table 5. Examining individual structures, the greatest post-intervention growth was observed for consonant+glide clusters, with simpler clusters of larger sonority distance improving the most. Growth in other classes of sound structures was either small, transient, or unattested in the monitored set.[2]

Effect size

To estimate an intervention effect size comparable with similar single-case experimental designs in clinical phonology, standard mean difference was calculated following Gierut

[2]It should be noted that the only singleton consonant at 0% accuracy across baselines for any of the children was trill /r/; thus, the absence of singleton change (excepting Jaime at 1 Month Post) reflects, specifically, unchanged accuracy of trill /r/.

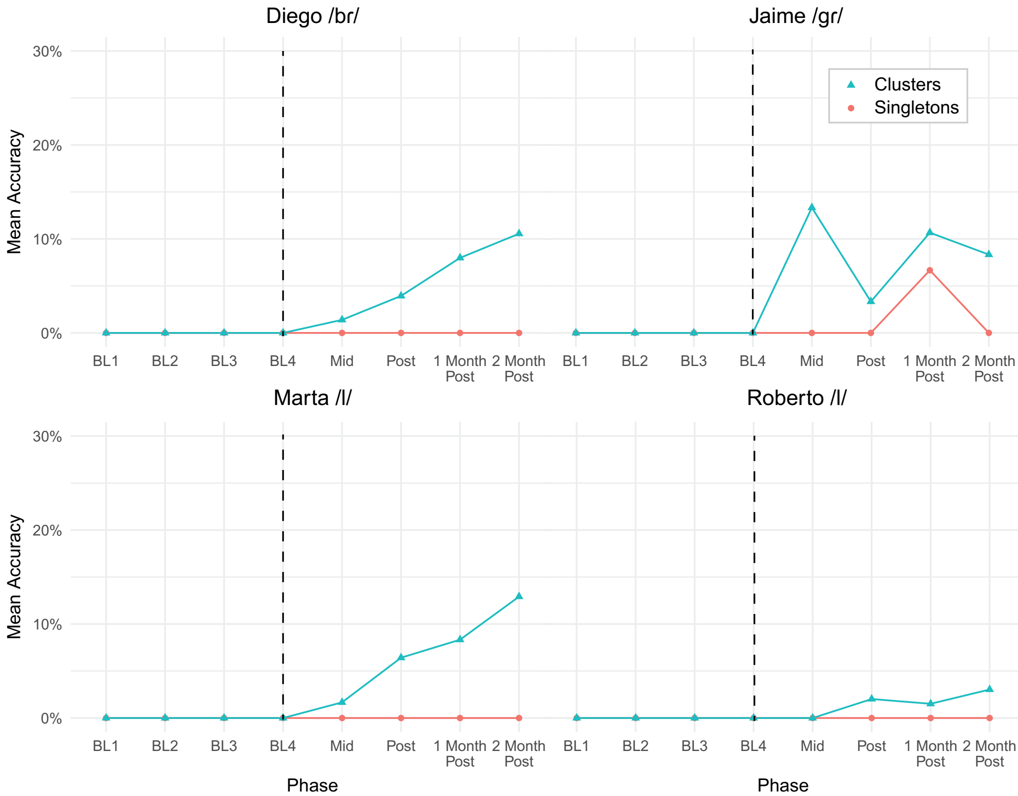

Figure 2. Composite accuracy of 0%-at-baseline singleton consonants and clusters. BL = baseline. Dashed lines indicate the onset of intervention.

Table 5. Accuracy change in monitored clusters.

Name	Cluster	Pre	Mid	Post	1 Month Post	2 Month Post
Diego	/kl/	0% (0)	0% (0)	20% (0.18)	0% (0)	0% (0)
/br/	/nj/	0% (0)	0% (0)	50% (0.35)	0% (0)	50% (0.35)
	/nw/	0% (0)	0% (0)	33% (0.27)	0% (0)	0% (0)
	/bj/	0% (0)	33% (0.27)	0% (0)	33% (0.27)	0% (0)
Jaime	/fr/	0% (0)	33% (0.27)	0% (0)	0% (0)	0% (0)
/gr/	/kr/	0% (0)	33% (0.27)	0% (0)	0% (0)	25% (0.22)
	/bw/	0% (0)	67% (0.27)	33% (0.27)	67% (0.27)	33% (0.27)
Marta	/kl/	0% (0)	0% (0)	20% (0.18)	0% (0)	0% (0)
/l/	/mw/	0% (0)	0% (0)	33% (0.27)	0% (0)	33% (0.27)
	/nj/	0% (0)	0% (0)	50% (0.35)	50% (0.35)	50% (0.35)
	/fj/	0% (0)	33% (0.27)	0% (0)	33% (0.27)	67% (0.27)
	/sj/	0% (0)	0% (0)	25% (0.22)	50% (0.25)	75% (0.22)
Roberto	/kl/	0% (0)	0% (0)	20% (0.18)	0% (0)	0% (0)
/l/	/pr/	0% (0)	0% (0)	33% (0.27)	0% (0)	0% (0)

Treatment targets are shown under participant names. Standard errors in parentheses.

et al. (2015), using accuracy of singletons and clusters produced with 0% accuracy at the pre-intervention assessment and the pooled standard deviation of accuracy of those structures across baselines (0.19). Standard mean difference permits comparison of the impact of intervention, across study participants, in a way that accounts for participants' pooled variance as well as individual accuracy change during baseline. Note that this metric does not exclude those 0%-accurate sounds that were produced with some accuracy during the baseline sessions (not exceeding ±10% change). This was necessary to observe each participant's variance in 0%-accurate sounds across baseline sessions and to determine the standard deviation of accuracy at baseline, which cannot be zero, as it is the denominator of the standard mean difference calculation.

The standard mean difference for each participant's growth in accuracy of their monitored sounds is shown in Table 6. Jaime demonstrated the largest intervention effect at 8.5, followed by Diego at 1.5, Roberto at 0.8, and Marta at 0.2. Also included in Table 6 is each participant's accuracy during the baseline period for those sounds that were 0% accurate at the pre-intervention session. Any accuracy improvement to those sounds during baseline is indicative of pre-intervention variability and potential for growth not attributable to the intervention and, thus, reduces the standard mean difference effect size. During baseline, Jaime and Marta showed greater variability in their monitored structures, whereas Diego and Roberto showed little to no change in accuracy of their monitored structures. Table 6 also displays each participant's PCC-R (Shriberg et al., 1997) for comparison with a typical clinical metric.

Inventories

Each child's phonetic and cluster inventories are shown at pre- and post-intervention timepoints in Table 7. At the post-session, Marta added [ɾ, dw, kw, mj, nj, and tj], Roberto added [bj, fj, kj, mj, nj, pj, and sj], and Jaime added [v, ɾ, ʧ, dj, fj, nw, pj, sj, tj, and tw]. Diego did not add any consonants or clusters at the post-session. Furthermore, Marta "lost" one non-ambient (i.e. not occurring in the target variety of Spanish) consonant [ʔ], Roberto lost three non-ambient consonants [l̃, ʃ, ʦ], and Jaime lost two non-ambient clusters [kᵊl, ɬ w] at their post sessions.

Intelligibility in context scale and parent report

The ICS (McLeod, 2020; McLeod et al., 2012) was completed by a parent of each child at pre- and post-intervention sessions. The ICS is a subjective intelligibility-rating screener, which asks the respondent to rate their child's intelligibility on a 5-point scale for different communication partners (i.e. parent, immediate and extended family, friends, acquaintances, teachers, and strangers). The average total score for each child increased post intervention. Roberto's score increased by 0.57 (pre = 2.57; post = 3.14), Marta's score increased by 0.58 (pre = 2.75; post = 3.33), Diego's score increased by 1.15 (pre = 2.14; post = 3.29), and Jaime's score increased by 0.79 (pre = 2.14; post = 2.93). Parents' ratings on the ICS were mostly

Table 6. Standard mean difference effect size and PCC-R.

Name	Target	Baseline	Intervention	SMD	PCC-R Pre	PCC-R Post
Diego	/bɾ/	0.3%	3.2%	1.5	57.2%	64.9%
Jaime	/gɾ/	3.8%	20.0%	8.5	54.2%	59.8%
Roberto	/l/	0.0%	1.5%	0.8	63.1%	67.1%
Marta	/l/	3.8%	4.2%	0.2	53.7%	60.1%

Following Gierut et al. (2015), standard mean difference was calculated using accuracy of singletons and clusters produced with 0% accuracy during the pre-intervention assessment and the standard deviation of accuracy in those sounds during baseline pooled across participants (0.19). Baseline = mean accuracy across baseline sessions. Intervention = mean accuracy across intervention and post sessions. SMD = standard mean difference.

Table 7. Phonetic and cluster inventories at pre- and post-sessions.

Name	Pre	Post
Diego /bɾ/	[p, b, t, d, k, ɡ, m, n, ɲ, f, v, s, ɹ̣, x, h, ʦ, ʧ, w, β, ɣ, ð, l, ɾ, j]	[p, b, t, d, k, ɡ, m, n, ɲ, f, s, ɹ̣, x, h, ʧ, w, β, ɣ, ð, l, j]
	[fw, sw, bw, pw, kw, tw]	[fw, sw, bw, pw, kw, tw]
Jaime /ɡɾ/	[p, b, t, tʰ, d, k, ɡ, m, n, ɲ, f, s, ɹ̣, x, h, ɬ, ʦ, w, β, ɣ, ð, l, j]	[p, b, t, tʰ, d, k, ɡ, m, n, ɲ, f, **v**, s, ɹ̣, x, ɬ, ʦ, ʧ, w, β, ɣ, ð, l, ɾ, j]
	[ɬ w, k l, mw, mj, fw, fl, pw, pl, kw, kj, kl]	[**nw**, mj, mw, fl, **sj**, fw, **fj, dj**, pw, pl, kl, **tj, tw**, kw, **pj**, kj]
Marta /l/	[p, b, t, d, k, ɡ, ʔ, m, n, ɲ, f, v, s, ɹ̣, x, h, w, β, ɣ, ð, l, j]	[p, b, t, d, k, ɡ, m, n, ɲ, f, s, ɹ̣, x, h, w, β, ɣ, ð, l, ɾ, j]
	[dj, bw]	[**nj, mj**, dj, **dw, kw, tj**]
Roberto /l/	[p, b, t, d, k, ɡ, m, n, ɲ, f, v, s, ʃ, ɹ̣, x, ʦ, ʧ, w, β, ɣ, ð, l, l, ɾ, j]	[p, b, t, d, k, ɡ, m, n, ɲ, f, v, s, ɹ̣, x, ʧ, w, β, ɣ, ð, l, ɾ, j]
	[dj, tj]	[**nj, mj, sj, fj, bj, kj, pj**]

Treatment targets are shown under participant names. Singletons or cluster added to the inventory in the post-intervention session are bolded.

consistent with their report on a post-intervention survey. All participants' parents indicated that they noticed improvement in their child's speech sound production and intelligibility, except for Diego's parent, who indicated that she did not notice change. However, she did report improvement during the 1-month follow-up session.

Discussion

In this study, four Spanish-dominant Spanish-English bilingual children with SSD participated in a speech intervention targeting complex phonological structures in Spanish. Some phonological growth was demonstrated by all the children following intervention; however, there was considerable variation in each child's learning of the targeted structure and generalization to untargeted structures. These findings may support the feasibility of this intervention for Spanish-speaking children; however, considerable individual variability limits the generalizability of these findings. Based on standard mean difference, the impact of targeting complex phonological structure on system-wide generalization was greater for participants who learned a complex cluster than those who learned a complex singleton, which aligns with the results of two case studies that have targeted clusters in Spanish (Anderson, 2002; Barlow, 2005). We thus begin to extend the efficacy of an intervention approach that has been shown to be appropriate for monolingual English-speaking children (Gierut, 1999; Gierut & Champion, 2001; Gierut et al., 1996) to Spanish-speaking bilingual children.

Individual differences

Although all participating children demonstrated some degree of phonological growth, there was great variability in their individual responses to the intervention. To facilitate a discussion of each child's results, we will summarize the findings for each participant before reflecting upon potential explanatory factors.

Diego's intervention target was the complex cluster /bɾ/. Given limited accuracy gains with his target, acquiring the targeted structure was more challenging for Diego than for the other participants, including Jaime, who was also trained with a similar complex cluster. Across the baseline period, his accuracy and speech production patterns were relatively stable. Despite achieving only 22% accuracy with the cluster target, his relative stability at baseline allows his post-intervention generalization to be more confidently attributed to the intervention, as reflected in a standard mean difference of 1.5, which is higher than the two children who learned singleton targets and lower than the child who demonstrated greater success in acquiring the cluster target.

Jaime was trained to produce the cluster target /gɾ/. In stark contrast to Diego, his production accuracy in monitored sounds at baseline was markedly variable. He was also the oldest participant in the study. Although fluctuating, his learning of the targeted cluster trended upward, peaking at 74% accuracy at session 12. Of all the children, he demonstrated the most phonological generalization during and immediately post-intervention with a standard mean difference of 8.5, even after modulating for his variance in accuracy during baseline.

Roberto's target was the complex singleton /l/. He demonstrated relatively delayed improvement on his treatment target, despite a singleton target being ostensibly easier to master than a cluster target (e.g. Rvachew & Nowak, 2001). Despite this, he ultimately reached a relatively high level of accuracy with his intervention target, surpassed only by Marta, who was also trained with a singleton target. In keeping with his delayed trajectory in learning the intervention target, he was also the only participant to demonstrate no generalized growth in 0%-at-baseline sounds at the midpoint of intervention. Like Diego, his variance in monitored accuracy during baseline was very small. He also demonstrated only limited phonological generalization attributable to the intervention, with a standard mean difference of 0.8.

Marta's target was also the complex singleton /l/. She was the youngest participant and the only one to reach the performance criterion for intervention completion, demonstrating 100% accuracy with her intervention target across three sessions. She was thus the most rapid and successful in learning the target. She also demonstrated relatively broad phonological growth; however, her standard mean difference was the lowest at 0.2. This may seem unexpected, but it is attributable to greater variance in her monitored production accuracy during baseline, which minimizes the amount of phonological generalization, which can be ascribed to the applied intervention.

Taken together, participating children who demonstrated phonological systems with more variability in production accuracy of 0%-accurate sounds prior to intervention were also those that demonstrated the greatest post-intervention phonological growth. After adjusting this growth for variability during baseline, only a small portion of the observed growth could be attributed to Marta's intervention with a singleton target. However, the large amount of broad phonological growth observed for Jaime, who had a cluster target, still amounted to a relatively larger effect size, despite attenuating for his variability. The children who were less variable in their productions did not appear to learn their targeted structures as efficiently as those who were more variable. This relationship between variability and readiness for phonological learning is discussed further as a theoretical implication.

Clinical and theoretical implications

One of the goals of this study was to describe patterns of generalization following intervention. Clinically, our understanding of the scope of expected generalization from particular linguistic structures trained in intervention can support ideal intervention target selection to maximize improvement in less time (e.g. Barlow & Enríquez, 2008).

Theoretically, speech acquisition and generalization patterns have been associated with universal implicational relationships and relative markedness among phonological structures (Gierut, 2007). It is thus critical that we document phonological generalization patterns to inform both our clinical and linguistic understanding of target structures, especially beyond monolingual English speakers.

One observation that spanned all participants' generalization patterns is the emergence or maturation of consonant clusters during or post intervention. Per cross-linguistic implicational laws (e.g. Watts & Rose, 2020), this is expected for Diego and Jaime with cluster targets, but unexpected for Marta and Roberto with singleton targets. The emergence of other clusters and singletons is expected following intervention targeting clusters (Gierut, 1999; Gierut & Champion, 2001), but intervention targeting singletons is only expected to stimulate across-class growth in singleton structures, at least in English (Gierut, 2007). This difference could be influenced by language-specific variables, methodology, or both.

Most of the observed growth in consonant clusters in this study was within consonant +glide clusters (e.g. /bw/, /fj/), whose status as "true" branching onset clusters in Spanish is debated (e.g. Harris, 1983; Senturia, 1998). Thus, the growth in glide clusters observed in this study may not be analogous to growth in true branching clusters. Additionally, because intervention targeting a singleton consonant is not expected to cause significant growth in branching onset clusters, the observed improvement to consonant+glide clusters may provide additional evidence for the status of these sequences as simpler singleton onsets followed by a nucleus constituent in these children's phonological systems. Nevertheless, participants trained with a singleton also demonstrated limited improvement to other clusters (i.e. /kl/ and /pɾ/) that are not subject to the same debate about their branching onset status (Harris, 1983). Thus, some, but not all the observed growth in clusters following intervention targeting a singleton, may be attributable to the constituency of glide clusters within the syllable.

It may also be the case that generalization following intervention targeting singleton consonants does not exclusively affect individual consonants. Most research establishing generalization patterns following intervention with singleton targets only reported singleton or general consonant generalization data, as consonant clusters were not expected to change and thus were not monitored (Dinnsen & Elbert, 1984; Elbert & McReynolds, 1985; Flint & Costello Ingham, 2005; Gierut, 1990, 1991; Gierut et al., 1987; Gierut & Morrisette, 2012; Gierut et al., 1996; Gierut & Neumann, 1992; Miccio et al., 1999; Powell et al., 1998; Rvachew & Nowak, 2001). For few studies in which complex singletons were targeted and consonant clusters were included in generalization results, most, if not all, evidenced some growth in untargeted consonant clusters (Elbert & McReynolds, 1979; Gierut, 1998a; Miccio & Ingrisano, 2000; Rvachew & Bernhardt, 2010). The observed generalized improvement in clusters for the two Spanish-speaking children in this study who learned a singleton target is thus aligned with similar results from intervention with monolingual English-speaking children.

The question that follows is how we might explain growth in phonological structures that would not necessarily be predicted by implicational relationships (i.e. singletons do not imply consonant clusters; Gierut, 2007; Watts & Rose, 2020). The observable improvement in sound structures that are more complex than the target itself could be analogous to the effect of a trigger within a dynamic system, which has been cited in explanation of the spontaneous emergence of complex motor behaviours (Fogel & Thelen, 1987; Thelen, 1995; Thelen & Smith, 2007). Within dynamic systems theory, the observation of broad or cascading improvements in speech-sound production following introduction of a new structure into the child's dynamic phonological system does not need to exclusively follow the patterns of linguistic universals or implicational laws (Rvachew & Bernhardt, 2010).

Of course, there is a large body of research that supports the role of implicational laws in predicting generalization patterns following speech intervention; thus, a comprehensive explanation must account for the roles of linguistically predictable generalization and the dynamic and somewhat unpredictable nature of a child's motoric and cognitive-linguistic systems. Children in this study trained with the most complex cluster targets demonstrated more phonological growth in similarly complex and implied less complex structures (i.e. clusters of similar and higher sonority distances) than children who were trained with singleton targets of lesser complexity. This is congruent with predictions based on implicational hierarchies (e.g. Gierut, 1999; Gierut & Champion, 2001). Nevertheless, some limited growth in true branching onset clusters was also observed in children trained with a singleton target, which indicates that not all growth is predictable from these linguistic relationships, but that the introduction of a new complex structure into a child's phonological system can also stimulate growth in as-yet unpredictable ways.

The dynamic nature of a child's development may also provide insight into the relationship between variability and readiness for phonological learning shown in these data. Much of the change that occurs within dynamic systems can occur covertly. Multiple small changes occurring across interacting systems may go unnoticed or appear as inconsistencies or variability. However, at some indeterminate point, a small change can trigger one or many observable behavioural changes, such as the emergence of a new, more complex sound production pattern. In this scenario, the new behaviour is not the product of the one small trigger change, but rather the product of accumulated changes prior to and including the trigger (Thelen & Smith, 2007). In other words, children who were demonstrating greater variability prior to and during intervention may have been in a state primed for phonological change, facilitated by the introduction of a complex structure into their phonological system.

Limitations

This study used a single-case experimental design with multiple baselines. It differed from a canonical multiple-baseline design in that baselines were not staggered across participants. Instead, baseline variance across participants was pooled and included in the standard mean difference effect size metric, which permits the pool of participants at baseline to collectively provide control against which the treatment effect may be observed. Other single-case designs, such as alternating ABAB designs, can offer greater experimental control by allowing a participant to serve, exclusively, as their own control; however, these designs may not be effective when the intervention effect is expected to endure after treatment, as is the case with this intervention (Byiers et al., 2012; McReynolds & Kearns, 1983). Nevertheless, there are limitations to the degree of experimental control that can be offered with this design, and both replication and expansion of this work are needed.

Both participants with singleton targets were trained on /l/. This consonant is considered less complex in Spanish relative to English (McLeod & Crowe, 2018); however, its relative complexity may be different in the unique phonological system of a bilingual child

(Fabiano-Smith & Goldstein, 2010). Although it was the least accurate, most complex singleton available for target selection in both cases, it may not be an ideal exemplar of a complex singleton target in Spanish. Other Spanish singletons, particularly /ʧ/, /ɾ/, and /r/ were excluded as potential targets due to positional constraints or the presence of complex gestures complicating their singleton status (Berns, 2013). However, these singletons could make for more complex intervention targets due to their later normative age of acquisition and articulatory complexity (Buchwald, 2017; Stokes & Surendran, 2005). Consequently, the very small effect sizes observed for the children with singleton targets may reflect the impact of moderately complex, rather than maximally complex, singleton targets.

Summary and future directions

The findings from this study, in concordance with earlier case studies (Anderson, 2002; Barlow, 2005), begin to extend the efficacy of targeting complex consonant clusters (i.e. /bɾ/ and /gɾ/) in intervention for Spanish-speaking bilingual children with SSD. However, broad phonological generalization was less apparent in children trained with a moderately complex singleton (i.e. /l/). Individual differences in responsiveness to the intervention also highlight the interconnected roles of variability, linguistic complexity, and the dynamic nature of a child's phonological system in speech intervention outcomes.

Given the initial nature of this study, replication and expansion are needed to understand the generalizability of these findings. Future investigations should consider examining a greater diversity of linguistic profiles, including monolingual and multilingual speakers of languages other than English. For instance, a study of speech intervention with monolingual Spanish-speaking children would permit better isolation of the role of Spanish phonology in optimal target selection. Related work may also benefit from more extensive monitoring for phonological generalization to better characterize broad phonological growth within and across languages. Finally, the role of baseline variability on the intervention outcome merits further investigation and may be an important consideration when examining related research questions in larger cohort studies or randomized controlled trials.

Acknowledgments

We express our gratitude to our young participants and their families. We thank the members of the Phonological Typologies Lab at San Diego State University, particularly Yureli Lopez, Mayerling Ovalles, and Katarina Richard for their leadership and coordination, as well as Jennifer Zaragoza, Monique Bareño, and Claudia Tellez for assistance with data collection. We are also grateful to Irina Potapova, Jennifer Taps Richard, and Leah Fabiano-Smith for their support, collaboration, and consultation in the development of this work.

Disclosure statement

The authors have no declarations of interest to report.

Funding

This work was supported by the National Institutes of Health under Grants NIDCD F31 DC017697 and NIDCD R21 DC01720.

ORCID

Philip Combiths http://orcid.org/0000-0003-3683-1650

References

Almost, D., & Rosenbaum, P. (1998). Effectiveness of speech intervention for phonological disorders: A randomized controlled trial. *Developmental Medicine and Child Neurology*, 40(5), 319–325. https://doi.org/10.1111/j.1469-8749.1998.tb15383.x

Anderson, R. T. (2002). Onset clusters and the sonority sequencing principle in Spanish: A treatment efficacy study. In F. Windsor, M. L. Kelly, & N. Hewlett (Eds.), *Investigations in clinical phonetics and linguistics*(pp. 213–224). Lawrence Erlbaum.

Baker, E., & Williams, A. L. (2010). Complexity approaches to intervention. In A. L. Williams, S. McLeod, & R. J. McCauley (Eds.), *Interventions for speech sound disorders in children* (pp. 95–116). Brooks Publishing.

Baker, E., Williams, A. L., McLeod, S., & McCauley, R. (2018). Elements of phonological interventions for children with speech sound disorders: The development of a taxonomy. *American Journal of Speech-Language Pathology*, 27(3), 906–935. https://doi.org/10.1044/2018_AJSLP-17-0127

Barberena, L. D., Mota, H. B., & Keske-Soares, M. (2015). As mudanças fonológicas obtidas pelo tratamento com o modelo ABAB-Retirada e Provas Múltiplas em diferentes gravidades do desvio fonológico [Phonological changes obtained by treatment based on ABAB-Withdrawal and Multiple-Probes Approach in different seve. *Revista CEFAC*, 17(1), 44–51. https://doi.org/10.1590/s1516-18462013005000053

Barlow, J. A. (2003). Asymmetries in the acquisition of consonant clusters in Spanish. *Canadian Journal of Linguistics*, 48(3–4), 179–210. https://doi.org/10.1017/S0008413100000645

Barlow, J. A. (2005). Phonological change and the representation of consonant clusters in Spanish: A case study. *Clinical Linguistics & Phonetics*, 19(8), 659–679. https://doi.org/10.1080/02699200412331279794

Barlow, J. A. (2012). *Little PEEP: Shorter protocol for the evaluation of English phonotactics*. Phonological Typologies Lab, San Diego State University.

Barlow, J. A., & Combiths, P. (2019). *Evaluación de la Fonología Española* [*Assessment of Spanish Phonology*]. Phonological Typologies Lab, San Diego State University.

Barlow, J. A., & Enríquez, M. (2008). Theoretical perspectives on speech sound disorders in bilingual children. *Perspectives on Language Learning and Education*, 14(2), 3. https://doi.org/10.1044/lle14.2.3

Bauman-Waengler, J. (2008). *Articulatory and phonological impairments: A clinical focus* (3rd ed.). Allyn & Bacon.

Berns, J. (2013). *Friction between phonetics and phonology: The status of affricates*. Netherlands Graduate School of Linguistics.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 206–244). Blackwell.

Buchwald, A. (2017). Complexity in articulatory and segmental levels of production. *Cognitive Neuropsychology*, 34(7–8), 488–492. https://doi.org/10.1080/02643294.2017.1422484

Byiers, B. J., Reichle, J., & Symons, F. J. (2012). Single-subject experimental design for evidence-based practice. *American Journal of Speech-Language Pathology*, 21(4), 397–414. https://doi.org/10.1044/1058-0360(2012/11-0036)

Cataño, L., Barlow, J. A., & Moyna, M. I. (2009). Phonetic inventory complexity in the phonological acquisition of Spanish: A retrospective, typological study. *Clinical Linguistics & Phonetics*, 23(6), 446–472. https://doi.org/10.1080/02699200902839818

Ceron, M. I., Pagliarin, K. C., & Keske-Soares, M. (2013). Advances in the treatment of children with phonological disorders. *International Archives of Otorhinolaryngology*, 17(2), 189–195. https://doi.org/10.7162/S1809-97772013000200012

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and physics of speech* (pp. 283–333). Cambridge University Press.

Combiths, P., Amberg, R., Hedlund, G., Rose, Y., & Barlow, J. A. (in press). Automated phonological analysis and treatment target selection using AutoPATT. *Clinical Linguistics & Phonetics*.

Combiths, P. N., Barlow, J. A., & Sanchez, E. (2019). Quantifying phonological knowledge in children with phonological disorder. *Clinical Linguistics & Phonetics*, 33(10–11), 885–898. https://doi.org/10.1080/02699206.2019.1584247

Cummings, A. E., & Barlow, J. A. (2011). A comparison of word lexicality in the treatment of speech sound disorders. *Clinical Linguistics & Phonetics*, 25(4), 265–286. https://doi.org/10.3109/02699206.2010.528822

Dinnsen, D. A., & Elbert, M. (1984). On the relationship between phonology and learning. In M. Elbert, D. A. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulating child* (ASHA Monographs No. 22) (pp. 59–68). ASHA.

Elbert, M., & McReynolds, L. V. (1979). Aspects of phonological acquisition during articulation training. *Journal of Speech and Hearing Disorders*, 44(4), 459–471. https://doi.org/10.1044/jshd.4404.459

Elbert, M., & McReynolds, L. V. (1985). The generalization hypothesis: Final consonant deletion. *Language and Speech*, 28(3), 281–294. https://doi.org/10.1177/002383098502800305

Elbert, M. F., Dinnsen, D. A., & Powell, T. W. (1984). On the prediction of phonologic generalization learning patterns. *Journal of Speech and Hearing Disorders*, 49(3), 309–317. https://doi.org/10.1044/jshd.4903.309

Ertmer, D. J., & Ertmer, P. A. (1998). Constructivist strategies in phonological intervention: Facilitating self-regulation for carryover. *Language, Speech, and Hearing Services in Schools*, 29(2), 67–75. https://doi.org/10.1044/0161-1461.2902.67

Fabiano-Smith, L., & Goldstein, B. A. (2010). Early-, middle-, and late-developing sounds in monolingual and bilingual children: An exploratory investigation. *American Journal of Speech-Language Pathology*, 19(1), 66–77. https://doi.org/10.1044/1058-0360(2009/08-0036)

Fernández Vítores, D. (2017). *El Español: Una Lengua Viva [Spanish: A Living Language]*. https://cvc.cervantes.es/lengua/espanol_lengua_viva/pdf/espanol_lengua_viva_2017.pdf

Flint, C. B., & Costello Ingham, J. (2005). Pretreatment stimulability and percentage of consonants correct as predictors of across-phoneme generalization. *Contemporary Issues in Communication Science and Disorders*, 32(Spring), 53–63. https://doi.org/10.1044/cicsd_32_S_53

Fogel, A., & Thelen, E. (1987). Development of early expressive and communicative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology*, 23(6), 747–761. https://doi.org/10.1037/0012-1649.23.6.747

Gierut, J. A. (1990). Differential learning of phonological oppositions. *Journal of Speech and Hearing Research*, 33(3), 540–549. https://doi.org/10.1044/jshr.3303.540

Gierut, J. A. (1991). Homonymy in phonological change. *Clinical Linguistics & Phonetics*, 5(2), 119–137. https://doi.org/10.3109/02699209108985509

Gierut, J. A. (1998a). Natural domains of cyclicity in phonological acquisition. *Clinical Linguistics & Phonetics*, 12(6), 481–499. https://doi.org/10.3109/02699209808985239

Gierut, J. A. (1998b). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research*, 41(1), S85–S100. https://doi.org/10.1044/jslhr.4101.s85

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research*, 42(3), 708–726. https://doi.org/10.1044/jslhr.4203.708

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology*, 16(1), 6–17. https://doi.org/10.1044/1058-0360(2007/003)

Gierut, J. A. (2008). Fundamentals of experimental design and treatment. In D. A. Dinnsen & J. A. Gierut (Eds.), *Optimality theory, phonological acquisition and disorders* (pp. 93–121). Equinox.
Gierut, J. A. (2015). *Experimental designs and protocols. Learnability Project Working Papers*. Indiana University. https://doi.org/10.5967/K8CC0XMD
Gierut, J. A., & Champion, A. H. (2001). Syllable onsets II: Three-element clusters in phonological treatment. *Journal of Speech, Language, and Hearing Research*, 44(4), 886–904. https://doi.org/10.1044/1092-4388(2001/071)
Gierut, J. A., Elbert, M., & Dinnsen, D. A. (1987). A functional analysis of phonological knowledge and generalization learning in misarticulating children. *Journal of Speech and Hearing Research*, 30 (4), 462–479. https://doi.org/10.1044/jshr.3004.432
Gierut, J. A., & Morrisette, M. L. (2012). Age of word acquisition effects in treatment of children with phonological delays. *Applied Psycholinguistics*, 33(1), 121–144. https://doi.org/10.1017/S0142716411000294
Gierut, J. A., Morrisette, M. L., & Dickinson, S. L. (2015). Effect size for single-subject design in phonological treatment. *Journal of Speech, Language, and Hearing Research*, 58(5), 1464. https://doi.org/10.1044/2015_JSLHR-S-14-0299
Gierut, J. A., Morrisette, M. L., Hughes, M. T., & Rowland, S. (1996). Phonological treatment efficacy and developmental norms. *Language, Speech, and Hearing Services in Schools*, 27(3), 215–230. https://doi.org/10.1044/0161-1461.2703.215
Gierut, J. A., & Neumann, H. J. (1992). Teaching and learning/θ/: A non-confound. *Clinical Linguistics & Phonetics*, 6(3), 191–200. https://doi.org/10.3109/02699209208985530
Glaspey, A. M., & Stoel-Gammon, C. (2005). Dynamic assessment in phonological disorders: The Scaffolding Scale of Stimulability. *Topics in Language Disorders*, 25(3), 220–230. https://doi.org/10.1097/00011363-200507000-00005
Harris, J. W. (1983). *Syllable structure and stress in Spanish*. MIT Press.
Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case design technical documentation*. What Works Clearinghouse.
Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell.
Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C. (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language & Communication Disorder*, 35(2), 165–188. https://doi.org/10.1080/136828200247133
Law, J., Garrett, Z., & Nye, C. (2004). The efficacy of treatment for children with developmental speech and language delay/disorder: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 47(4), 924–943. https://doi.org/10.1044/1092-4388(2004/069)
McLeod, S. (2020). Intelligibility in context scale: Cross-linguistic use, validity, and reliability. *Speech, Language and Hearing*, 23(1), 9–16. https://doi.org/10.1080/2050571X.2020.1718837
McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology*, 27(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100
McLeod, S., Harrison, L. J., & McCormack, J. (2012). The intelligibility in context scale: Validity and reliability of a subjective rating measure. *Journal of Speech, Language, and Hearing Research*, 55(2), 648–656. https://doi.org/10.1044/1092-4388(2011/10-0130)
McReynolds, L. V., & Kearns, K. P. (1983). *Single-subject experimental designs in communicative disorders*. University Park Press.
Miccio, A. W., Elbert, M., & Forrest, K. (1999). The relationship between stimulability and phonological acquisition in children with normally developing and disordered phonologies. *American Journal of Speech-Language Pathology*, 8(4), 347–363. https://doi.org/10.1044/1058-0360.0804.347
Miccio, A. W., & Ingrisano, D. R. (2000). The acquisition of fricatives and affricates: Evidence from a disordered phonological system. *American Journal of Speech-Language Pathology*, 9(3), 214–229. https://doi.org/10.1044/1058-0360.0903.214
Mota, H. B., Keske-Soares, M., Bagetti, T., Ceron, M. I., & Filha, M. D. G. D. C. M. (2007). Análise comparativa da eficiência de três diferentes modelos de terapia fonológica [Comparative analyses

of the effectiveness of three different phonological therapy models]. *Pró-fono Revista De Atualização Científica*, 19(1), 67–74. https://doi.org/10.1590/S0104-56872007000100008

Pagliarin, K. C., Mota, H. B., & Keske-Soares, M. (2009). Análise da eficácia terapêutica em três modelos fonológicos de abordagem contrastiva [Therapeutic efficacy analysis of three contrastive approach phonological models]. *Pro-fono: Revista De Atualizacao Cientifica*, 21(4), 297–302. https://doi.org/10.1590/S0104-56872009000400006

Parker, S. G. (2002). *Quantifying the sonority hierarchy* [Doctoral dissertation]. GLSA, University of Massachusetts.

Pereira, L. F., & Mota, H. B. (2002). Tratamento fonológico baseado nos contrastes de oposições máximas [Phonological treatment based on maximum opposition contrasts]. *Pró-Fono Revista de Atualização Científica*, 14(2), 165–174.

Pinker, S. (1984). *Language learnability and language development*. Harvard University Press.

Powell, T. W. (1993). Phonetic inventory constraints in young children: Factors affecting acquisition patterns during treatment. *Clinical Linguistics & Phonetics*, 7(1), 45–57. https://doi.org/10.3109/02699209308985543

Powell, T. W., Elbert, M., & Dinnsen, D. A. (1991). Stimulability as a factor in the phonological generalization of misarticulating preschool children. *Journal of Speech and Hearing Research*, 34 (6), 1318–1328. https://doi.org/10.1044/jshr.3406.1318

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(2), 127–146. https://doi.org/10.3109/02699209808985217

Real Academia Española. (2011) . *Nueva gramática de la lengua española: Fonética y fonología*. Espasa Libros.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research*, 41(6), 1398–1411. https://doi.org/10.1044/jslhr.4106.1398

Roid, G. H., & Miller, L. J. (1997). *Leiter international performance scale-revised*. Stoelting Co.

Rose, Y., & Hedlund, G. (2017). *Phon (Version 2.2)* [Computer software].

Rvachew, S., & Bernhardt, B. M. (2010). Clinical implications of dynamic systems theory for phonological development. *American Journal of Speech Language Pathology*, 19(1), 34–50. https://doi.org/10.1044/1058-0360(2009/08-0047)

Rvachew, S., & Nowak, M. (2001). The effect of target-selection strategy on phonological learning. *Journal of Speech, Language, and Hearing Research*, 44(3), 610–623. https://doi.org/10.1044/1092-4388(2001/050)

Secord, W. A. (2007). *Eliciting sounds: Techniques and strategies for clinicians* (2nd ed.). Singular.

Senturia, M. B. (1998). *A prosodic theory of hiatus resolution* [Doctoral dissertation]. University of California.

Shriberg, L., Austin, D., Lewis, B. A., McSweeny, J., & Wilson, D. (1997). The Percentage of Consonants Correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, 40(4), 708–722. https://doi.org/10.1044/jslhr.4004.708

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders II: A conceptual framework for management. *Journal of Speech and Hearing Disorders*, 47(3), 242–256. https://doi.org/10.1044/jshd.4703.242

Shriberg, L. D., & Kwiatkowski, J. (1990). Self-monitoring and generalization in preschool speech-delayed children. *Language, Speech, and Hearing Services in Schools*, 21(3), 157–170. https://doi.org/10.1044/0161-1461.2103.157

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow transcription. *Clinical Linguistics & Phonetics*, 5(3), 225–279. https://doi.org/10.3109/02699209108986113

Smit, A. B., Mann Brumbaugh, K., Weltsch, B., & Hilgers, M. (2018). Treatment of phonological disorder: A feasibility study with focus on outcome measures. *American Journal of Speech-Language Pathology*, 27(2), 536–552. https://doi.org/10.1044/2017_AJSLP-16-0225

Sommers, R. K., Leiss, R. H., Delp, M. A., Gerber, A. J., Fundrella, D., Smith, R. M., Revucky, M. V., Ellis, D., & Haley, V. A. (1967). Factors related to the effectiveness of articulation therapy for kindergarten, first, and second grade children. *Journal of Speech and Hearing Research*, 10(3), 428–437. https://doi.org/10.1044/jshr.1003.428

Stokes, S. F., & Surendran, D. (2005). Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *Journal of Speech, Language, and Hearing Research*, 48(3), 577–591. https://doi.org/10.1044/1092-4388(2005/040)

Storkel, H. L. (2018). The complexity approach to phonological treatment: How to select treatment targets. *Language, Speech, and Hearing Services in Schools*, 49(3), 1–19. https://doi.org/10.1044/2017_LSHSS-17-0082

Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist*, 51(2), 79–95. https://doi.org/10.1037/0003-066X.50.2.79

Thelen, E., & Smith, L. B. (2007). Dynamic systems theories. In W. Damon & R. M. Lemer (Eds.), *Handbook of child psychology: Theoretical models of human development* (5th ed., pp. 258–312). John Wiley & Sons.

Thompson, C. K., Shapiro, L. P., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The Complexity Account of Treatment Efficacy (CATE). *Journal of Speech, Language, and Hearing Research*, 46(3), 591–607. https://doi.org/10.1044/1092-4388(2003/047)

Van Horne, A. J. O., Fey, M., & Curran, M. (2017). Do the hard things first: A randomized controlled trial testing the effects of exemplar selection on generalization following therapy for grammatical morphology. *Journal of Speech, Language, and Hearing Research*, 60(9), 2569. https://doi.org/10.1044/2017_JSLHR-L-17-0001

Watts, E., & Rose, Y. (2020). Markedness and implicational relationships in phonological development: A cross-linguistic investigation. *International Journal of Speech-language Pathology*, *22*(6), 1–14. https://doi.org/10.1080/17549507.2020.1842906

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2012). *Preschool language scales fifth edition Spanish*. Pearson.