Title
Advancing efficient and equitable intervention for children with phonological disorder

Author
Combiths, Philip Nichol

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Advancing efficient and equitable intervention for children with phonological disorder

A dissertation submitted in partial satisfaction of the requirements
for the degree Doctor of Philosophy

in

Language and Communicative Disorders

by

Philip Combiths

Committee in charge:

San Diego State University

Professor Jessica Barlow, Chair
Professor Sonja Pruitt-Lord, Co-Chair
Professor Gregory Keating
Professor Giang Pham

University of California San Diego

Professor Seana Coulson
Professor Marc Garellek

2021

The Dissertation of Philip Combiths is approved, and is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California San Diego

San Diego State University

2021

DEDICATION

Dedicated to Maria Elena Guerena “Meg” Combiths.

EPIGRAPH

My humanity is bound up in yours, for we can only be human together.

– Desmond Tutu

TABLE OF CONTENTS

Dissertation Approval Page iii

Dedication iv

Epigraph v

Table of Contents vi

List of Figures xii

List of Tables xiii

Acknowledgements xv

Vita xix

Abstract of the Dissertation xxi

Chapter 1: Introduction 1

Phonology and Phonological Disorder 2

Efficient Intervention for Phonological Disorder 4

Perspectives on Treatment Efficiency 6

Traditional Approach 7

Complexity Approach 11

Comparative Evidence 17

Equitable Intervention for Phonological Disorder 24

Overview of Studies 26

References 28

Chapter 2: Automated Phonological Analysis and Treatment Target Selection Using AutoPATT 38

Abstract 39

Introduction 40

The Current Study 45

Method 46

Participants and Transcriptions 46
Automated and Manual Data 47
Analyses 48
Results 49
Agreement 49
Accuracy 49
Qualitative Results 50
Discussion 53
Clinical Implications 54
Limitations and Future Directions 55
Conclusion 57
Chapter 3: Quantifying Phonological Knowledge in Children with Phonological Disorder 65
Abstract 66
Introduction 66
Phonemic Inventory Analysis 67
Production Accuracy 68
Current Study 69
Method 69
Participants 69
Data Transformation 70
Variables 71
Analyses 71
Results 72
Descriptive Statistics 72
Logistic Regression 72
Receiver Operating Characteristic Curve Analysis 72

Discussion....74
Implications for Assessment....75
Limitations and Future Directions....76
References....77
Chapter 4: Influences of Phonological Context on Tense Marking in Spanish–English Dual Language Learners....81
Abstract....82
Introduction....82
Word-Final Tense Morphology in Linguistically Diverse Populations....82
Influences on Tense-Morpheme Production....83
PT Marking and Word-Final Phonological Context....84
3s Marking and Expansion of Phonological Context....86
The Current Study....86
Method....87
Participants....87
Procedure....87
Language-Sample Measures....87
Phonological Transcription....88
Morpheme Marking....88
Phonological Predictor Variables....88
Lexical and Sublexical Control Variables....90
Coding Procedures....90
Statistical Analysis....90
Results....91
Number of Surrounding Consonants....91
Preceding C/V Context....91

Preceding Sonority Distance 92
Phonological Predictors of Tense-Morpheme Marking....................92
Differentiating Patterns of Tense-Morpheme Marking....................93
Discussion....................93
Modulating Word-Final Complexity....................95
Limitations and Future Directions....................95
References....................97
Chapter 5: Treatment Targets for Co-Occurring Speech-Language Impairment: A Case Study....................101
Abstract....................102
Introduction....................102
Overlapping Versus Co-Occurring Disorders....................103
Cross-Domain Treatment: Approaches. Techniques, and Targets....................103
Treatment Target Complexity and Generalization....................104
Complex Treatment Targets for PD-LD....................104
A Case Study....................105
Method....................106
Participant....................106
Generalization Probes....................106
Treatment Target and Stimuli....................107
Treatment Protocol....................107
Outcome Measures....................108
Results....................108
Baseline Stability....................108
Treatment Progress....................109
Phonology....................109

Grammatical Morphology....................110
Treatment Impact and Effect Size....................110
Discussion....................111
Generalization in the Phonological Domain....................111
Cross-Domain Generalization....................112
Limitations....................113
Clinical Implications....................113
References....................114
Appendix....................116
Chapter 6: Phonological Complexity in Intervention for Spanish-Speaking Children with Speech Sound Disorder....................120
Abstract....................121
Introduction....................122
The Current Study....................125
Method....................125
Participants....................125
Assessment Procedures....................128
Dependent and Descriptive Variables....................129
Experimental Design....................131
Intervention Targets and Materials....................132
Intervention Protocol....................133
Transcription and Analysis....................134
Results....................134
Intervention and Target Accuracy....................135
Generalization to Monitored Sounds....................136
Effect Size....................139

Inventories 140
Intelligibility in Context Scale and Parent Report 141
Discussion 142
Individual Differences 142
Clinical and Theoretical Implications 144
Limitations 147
Summary and Future Directions 148
References 149
Chapter 7: General Discussion 157
References 164

LIST OF FIGURES

Figure 3-1. Phonemic inventory classification. 73

Figure 3-2. ROC curve for phoneme classification via consonant accuracy by sample type...74

Figure 4-1. Syllable structure (A–D), sonority distance (E–G), and resyllabification (H–I) in contexts of varied word-final complexity. 85

Figure 4-2. Findings for marking rates of past tense and third person singular. For number of surrounding consonants and word-final sonority distance, low resyllabification = morpheme resyllabification at 0.00, and high resyllabification = morpheme resyllabification at 0.01. For preceding consonant/vowel (C/V) context, Low resyllab. and High resyllab. groups represent a median split of the morpheme-resyllabification values (Low resyllab. ≤ 0.0005, High resyllab. > 0.0005). Error bars indicate 95% confidence intervals. 92

Figure 4-3. Marking rate by preceding consonant/vowel (C/V) context for past tense (PT) and third person singular (3s) morphemes and their corresponding nonmorphemic homophones (/d/, /t/; /z/, /s/) in word-final position. Error bars indicate 95% confidence intervals. 94

Figure 5-1. Sample target production rating scale. 108

Figure 5-2. Production accuracy rating (1–7) of /–lps/ cluster production in target verbs by week. 109

Figure 5-3. Composite accuracy of 0%-at-baseline singletons and clusters by word position. 109

Figure 6-1. Intervention target accuracy. 136

Figure 6-2. Composite accuracy of 0%-at-baseline singleton consonants and clusters. 138

LIST OF TABLES

Table 1-1. Implicational laws for complex treatment target selection. 12

Table 1-2. Non-exhaustive summary of studies comparing target mastery following treatment of simple versus complex sound structures. 19

Table 1-3. Non-exhaustive summary of studies comparing generalization following treatment of simple versus complex sound structures. 22

Table 2-1. Interrater reliability for AutoPATT and manual analyses as percent agreement, Cohen’s kappa, and Scott’s pi. 50

Table 2-2. AutoPATT analysis reliability as percent correct, Cohen’s kappa, and Scott’s pi. 50

Table 2-3. Manual analysis reliability as percent correct, Cohen’s kappa, and Scott’s pi. 50

Table 2-4. Unique inventory errors. 52

Table 3-1. Mean consonant accuracy by phonemic inventory, age of acquisition, and participant age. 72

Table 3-2. Potential consonant accuracy cutoff values predictive of phoneme status. 73

Table 3-3. Phoneme classification metrics for potential accuracy cutoff values by sample type. 74

Table 4-1. Abbreviated findings for word-final phonological-context effects on regular marking of past tense and third person singular. 84

Table 4-2. Contextual examples of past tense and third person singular tokens coded as marked or unmarked. 89

Table 4-3. Regression variables. 90

Table 4-4. Model comparisons for measures of phonological context. 93

Table 4-5. Summary of findings for phonological influences on marking rates for past tense and third person singular. 94

Table 5-1. Baseline percent correct singletons and clusters by position, percentage of consonants correct–revised (PCC-R), and monosegmental regular grammatical morpheme marking rate. 106

Table 5-2. Baseline phonetic inventory of singletons and clusters, by word position. 107

Table 5-3. Schedule of baseline/generalization probes, periods of treatment, and breaks during which no treatment occurred. 107

Table 5-4. Percentage of consonants correct-revised (PCC-R) and regular word-final morpheme marking rate stability across baselines. 109

Table 5-5. Zero-percent-at-baseline word-initial singletons and clusters showing change over the course of treatment. 110

Table 5-6. Zero-percent-at-baseline word-final singletons and clusters showing change over the course of treatment. 110

Table 5-7. Regular word-final morpheme marking rates. 110

Table 5-8. Morpheme marking rates from pre- and posttreatment language samples. 111

Table 5-9. Percentage of consonants correct-revised (PCC-R) by time point with logistic regression results. 111

Table 6-1. Participant characteristics at initial assessment. 127

Table 6-2. Inclusionary and exclusionary criteria. 127

Table 6-3. Assessment measures by study phase. 129

Table 6-4. 0%-at-baseline sound structures monitored for generalization. 130

Table 6-5. Accuracy change in monitored clusters. 137

Table 6-6. Standard mean difference effect size and PCC-R. 140

Table 6-7. Phonetic and cluster inventories at pre and post sessions. 141

ACKNOWLEDGEMENTS

With all my heart, I thank my two advisors, Drs. Jessica Barlow and Sonja Pruitt-Lord. I could not have had a better pair of academic parents. Jessica, you were the first person to ever tell me that you believed I could make it here, and I would not be in this position if you had not asked me what I wanted to do with my life those several years ago. As a role model to me, you deftly exemplified both precision and deep, thoughtful compassion. Sonja, I have never met anyone with your incredible ability to guide others to become their best selves. You have what seems to be limitless compassion and generosity for others, and I believe anyone who crosses paths with you is lucky to do so. Together you both guided me in every aspect of this journey, from the esoteric and academic to the deeply personal. Perhaps the most incredible aspect of your co-mentorship is that you didn’t just set me up for success as a scientist—you also gave me the freedom to be a dreamer. You gave me the tools to grow, but you never dictated the direction. I will always treasure the care you have given me.

Mentorship comes in many forms, and I am grateful to all those who have guided me at different stages in this journey. I thank Dr. Jörg Matt, who taught me to love my plants and to love my stats, Jennifer Taps Richard and Charlotte Lopes, who taught me what it looks like to be an excellent and compassionate clinician, Dr. Leah Fabiano-Smith, who taught me to recognize my academic privilege and use that knowledge to uplift others, Dr. Eric Baković, who showed me how to make meaningful connections with academic peers, and Dr. Tracy Love, who fought so many battles for me and all the students in this program. I am so very grateful to the entire faculty of the SDSU/UCSD Joint Doctoral Program in Language and Communicative Disorders (JDP-LCD) for their advice, their kindness, and generosity with their time. I am fortunate to have a fantastic committee of experts providing guidance on this dissertation, and I thank Drs. Seana Coulson, Sarah Creel, Marc Garellek, Greg Keating, and Giang Pham for being so generous with your time and expertise.

To Irina, Nat, Brittany, Gabby, Erin, Jonathan, Steph, Niloo, and the rest of the JDP-LCD I could not have completed this journey without your kindness and support. This journey together has been unforgettable, and I treasure having shared it with you. Thank you to the past and current members of the Phonological Typologies Lab, especially Makayla Davis, Emilie Sanchez, Alexis Buck, Yureli Lopez, Mayerling Ovalles, and Katarina Richard, who worked closely with me on these projects. I thank my friends and family, who supported all the other aspects of life that made this dissertation possible. To Ian, who stuck it out with me through every disappointment and every success, I thank you for taking care of so many details and supporting me in my academic goals, even when they meant late nights, early mornings, and moving across the country. You have been my rock, keeping me grounded and helping me to take care of myself and take care of us.

Finally, a special thank you to my mother, Meg, to whom I dedicate this dissertation. You were and always will be my biggest champion. From my earliest memories, you instilled a penchant for curiosity and a love of learning. You always knew how to find common ground with others and had a soft spot in your heart for the unrecognized and underserved. I dedicate this work to your uplifting spirit in hopes that it makes a tiny contribution to a more equitable world for children with communication impairments and their families, in your honor.

Chapter 1 includes material from the dissertation author's unpublished integrative paper for the JDP-LCD, Combiths, P. (2018). *Treatment Target Selection for Phonological Intervention*. The dissertation author was the sole author of this paper.

Chapter 2, in full, is a reprint of the manuscript, 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.* The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of material as it appears in Combiths, P., Barlow, J., & Sanchez, E., (2019). Quantifying phonological knowledge in children with phonological disorder. *Clinical Linguistics & Phonetics, 33*(10–11), 885–898. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, is a reprint of material as it appears in Combiths, P., Barlow, J. A., Potapova, I., & Pruitt-Lord, S. (2017). Influences of phonological context on tense marking in Spanish–English dual language learners. *Journal of Speech, Language, and Hearing Research, 60*(8), 2199–2216. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, is a reprint of material as it appears in Combiths, P., Barlow, J.A., Richard, J.T., & Pruitt-Lord, S.L. (2019). Treatment targets for co-occurring speech-language impairment: A case study. *Perspectives of the ASHA Special Interest Groups, 4*(2), 240–256. The dissertation author was the primary investigator and author of this paper.

Chapter 6, in full, is a reprint of the manuscript, Combiths, P., Pruitt-Lord, S., Escobedo, A., & Barlow J.A. (submitted). Phonological complexity in intervention for Spanish-speaking children with speech sound disorder. The dissertation author was the primary investigator and author of this paper.

I also acknowledge the funding that I have been so privileged to receive in support of my research training and the studies included in this dissertation: an NIH Institutional Training Grant, “Neurocognitive Approaches to Communication Disorders” (T32DC007361; PI: Shapiro); the SDSU Lipinsky Fellowship; the Paul Pfaff Memorial Scholarship; the Dorothy Baronofsky Memorial Scholarship; a graduate assistantship with the SDSU Sage Project and Center for Regional Sustainability; and an NIH Predoctoral Training Grant, “Speech Sound Disorder in Young Bilingual Children” (F31DC017697; PI: Combiths). I would also like to recognize additional funding that supported research included in this dissertation: NIH R21DC017201 (PI:

Barlow); NIH R03DC012141 (PI: Pruitt-Lord); a Price Philanthropies research grant awarded to Sonja Pruitt-Lord; and the Summer Undergraduate Research Program at San Diego State University awarded to Jessica Barlow and Mayerling Ovalles.

VITA

EDUCATION

2021	Ph.D., Language and Communicative Disorders San Diego State University / University of California San Diego
2012	M.A., Investigation in Spanish Linguistics Universidad Complutense, Madrid
2009	B.A., Linguistics, Spanish University of California, Berkeley

RESEARCH FUNDING

2019-	R21 DC017201, Manipulating Linguistic Complexity to Improve Child Language Treatment Outcomes National Institutes of Health National Institute on Deafness and Other Communication Disorders Role: Research Associate
2019-2021	F31 DC017697, Speech Sound Disorder in Young Bilingual Children National Institutes of Health National Institute on Deafness and Other Communication Disorders Role: Principal Investigator
2015–2016 2017–2018	T32 DC007361, Neurocognitive Approaches to Communication Disorders National Institutes of Health National Institute on Deafness and Other Communication Disorders Role: Trainee
2016-2017	Lipinsky Fellowship in Language and Communicative Disorders San Diego State University Role: Fellow

PUBLICATIONS (PEER REVIEWED)

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., Barlow, J., & Sanchez, E., (2019). Quantifying phonological knowledge in children with phonological disorder. *Clinical Linguistics & Phonetics, 33*(10–11), 885–898.

Combiths, P., Barlow, J.A., Richard, J.T., & Pruitt-Lord, S.L. (2019). Treatment targets for co-occurring speech-language impairment: A case study. *Perspectives of the ASHA Special Interest Groups, 4*(2), 240–256.

Barlow, J.A., & Combiths, P. (2019). Monolingual speech production in a bilingual context. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (eds.), *Proceedings of the 19th International Congress of the Phonetic Sciences* (pp. 2305–2309). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Potapova, I., Kelly, S., Combiths, P., & Pruitt-Lord, S. (2018). Evaluating English morpheme accuracy, diversity, and productivity measures in language samples of developing bilinguals. *Language, Speech, and Hearing Services in Schools, 24(2), 260–276.*

Combiths, P., Barlow, J. A., Potapova, I., & Pruitt-Lord, S. (2017). Influences of phonological context on tense marking in Spanish–English dual language learners. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2199–2216.

PUBLICATIONS (SUBMITTED)

Combiths, P., Pruitt-Lord, S., Escobedo, A., & Barlow J.A. (submitted). Phonological complexity in intervention for Spanish-speaking children with speech sound disorder.

OTHER PUBLICATIONS

Combiths, P. (2012). *La fuerza ilocutiva de las optativas independientes en español: Sobre la optatividad autónoma de la partícula ojalá, la semántica del modo subjuntivo, y la fuerza ilocutiva emotiva de valoración* [Illocutionary force in independent optative constructions in Spanish: On the self-governing optativity of the particle ojalá, the semantics of the subjunctive mood, and the expressive illocutionary force of evaluation]. Master's thesis, Universidad Complutense, Madrid. E-Prints Complutense.

ACADEMIC AWARDS

2021	Graduate Student of the Year, College of Health and Human Services, SDSU
2020	Faculty Forward Award, SDSU
2020/2016	Meritorious Poster Submission, American Speech-Language and Hearing Association Convention
2020	Student Success / Instructionally Related Programs Award Money Matters: Grant Writing Workshops for Graduate Students
2019	Student Success / Instructionally Related Programs Award Storytelling for Science
2018/2019	NIH Travel Award, Symposium on Research in Child Language Disorders
2017	Travel Award, International Association for the Study of Child Language
2017	Paul Pfaff Memorial Scholarship
2016	Lipinsky Fellowship in Language and Communicative Disorders
2015/2017	Travel Award, SDSU Center for Clinical and Cognitive Neuroscience
2014	Dorothy Baronofsky Memorial Scholarship

ABSTRACT OF THE DISSERTATION

Advancing efficient and equitable intervention for children with phonological disorder

by

Philip Combiths

Doctor of Philosophy in Language and Communicative Disorders

San Diego State University, 2021

University of California San Diego, 2021

Jessica Barlow, Chair

Sonja Pruitt-Lord, Co-Chair

Phonological disorder is a language impairment with no known cause that primarily affects the phonological domain. It is a highly prevalent and impactful communication disorder, yet it has a relatively stagnant history of standardized or relational assessment techniques and traditional, bottom-up remediation approaches designed, primarily, for monolingual speakers of majority varieties of English. Although assessment and treatment for phonological disorder is effective, traditional approaches are not optimized for maximizing efficient phonological growth,

which may be achievable by incorporating knowledge of complex phonological structure into the target selection process. Further, the evidence base fails to address the diverse impairment profiles and language backgrounds of many children with phonological disorder. By drawing on our understanding of phonology, phonological analysis, and phonological development, we can work towards advancing efficient and equitable assessment and intervention approaches for a clinically and linguistically diverse population of children with phonological disorder.

Chapter 1 provides an overview of the current state of intervention for phonological disorder, highlighting areas of need related to the efficiency and equity of clinical services for this population. Chapter 2 describes a new assessment tool, AutoPATT, and a study examining its validity and accuracy for generating independent, descriptive phonological assessment measures. Chapter 3 investigates the relationship between phonemic inventories generated by AutoPATT and more traditional accuracy measures by comparing both measures derived from single-word productions of 275 English-speaking children with phonological disorder. Chapter 4 examines the phonological influences affecting explicit marking of English tense and agreement morphemes in the connected speech of typically developing Spanish-English bilingual children. Chapter 5 presents a case study exploring the use of a word-final treatment target containing complex phonological and morphological components to remediate co-occurring phonological and morphosyntactic deficits. Chapter 6 is a study comparing the efficacy and efficiency of treatment with complex consonant clusters and singleton targets in Spanish for Spanish-English bilingual children with phonological disorder. Chapter 7 concludes with a general discussion of the findings from these studies and describes pathways for continued work to advance the efficiency and equity of assessment and treatment for phonological disorder.

CHAPTER 1:

Introduction

Phonology and Phonological Disorder

Developmental language impairments, which can negatively affect the time course and ultimate attainment of language (Bishop et al., 2016; Gierut, 1998b) are highly prevalent communication disorders that collectively affect 8% or more of preschool and early school-age children (Black et al., 2015a; Eadie et al., 2014; Law et al., 2000; Shriberg et al., 1999; Tomblin et al., 1997; Wren et al., 2016). In this discussion, we are concerned with language impairments insofar as they impact, primarily, the phonological domain of language. When the acquisition or production of phonological structures is functionally impaired, such that a child's intelligibility or communication ability is significantly reduced compared to their typically developing peers, the impairment is most frequently classified as phonological disorder (PD; Shriberg & Kwiatkowski, 1982). When a developmental language impairment is similarly evidenced at the level of morphology or larger linguistic units (e.g., words or phrases) in the absence of another known the cause, the impairment has most recently been termed developmental language disorder (DLD; Bishop et al., 2017), although decades of prior research have used the term specific language impairment (e.g., Rice & Wexler, 1996). In either case, the etiology of these language impairments is poorly understood (Gierut, 1998b; Leonard, 2014), and their impact is apparent in both linguistic and non-linguistic domains (Clark & Lum, 2017; Dodd et al., 1989; Lewis et al., 2006).

In speech-language pathology, as the name itself suggests, there has traditionally been a divide between the clinical domains of speech (i.e., articulation and speech sound production) and language (i.e., grammar and rule-based use of linguistic units). However, the linguistic domain of phonology blatantly straddles this clinical division between speech and language. Human language evidences rule-based phonological structure that governs the contrastive units which combine to form morphemes and words. This is true in both spoken and sign languages (Brentari et al., 2018) and is thus not relegated only to the domain of speech. In spoken languages, this contrastive unit is the phoneme, an abstract and categorical representation of a speech sound unit that users of the same language variety perceive and produce according to mutually consistent parameters to differentiate word forms and create the basis of a shared, symbolic, and mutually intelligible language system.

When phonological representations are relatively consistent across individuals, as is the case with speakers of the same language variety, a sequence of phonemes, like /bæt/ in English, is consistently associated with one or a set of potential meanings, which can be further specified by context. Together, the phonological form and its associated semantic value(s) create a morpheme, lexeme, or word—in this case, “bat.” An English speaker can thus produce this word with confidence that another proficient English speaker would unambiguously perceive and interpret /bæt/ as the intended word “bat,” with its corresponding, context-bound meaning. The power of phonology and the phoneme unit is exemplified by the ability to transform one word into a completely unrelated word by substituting just a single phoneme. For instance, in English, /bæt/ “bat” becomes /pæt/ “pat” by substituting only the voiced bilabial plosive /b/ with its voiceless counterpart /p/. Acoustically, /b/ and /p/ are nearly identical, save for a difference in timing of the onset of voicing relative to the segment’s burst release. Such a small acoustic difference results in two words that are, semantically, completely distinct. For such a small acoustic difference to result in the production and/or perception of distinct words, the parameters that define the shared set of phoneme units in mutually intelligible language varieties must be calibrated across users of that language variety.

In language acquisition, the precise and timely development of one or more phonological systems is thus crucial to the progression of all aspects of language and is simultaneously no small feat. As illustrated above, the articulatory and perceptual parameters defining contrastive phonemes must be acquired. The acquisition and continued refinement of these parameters is, by itself, an impressive task. To give an outstanding example, Taa (also known as !Xóõ) is one language documented to contrast between 89 and 107 phonemes (for comparison, English and Spanish contrast between 38–50 and 22–25 phonemes, respectively; Ladefoged & Maddieson, 1996). Furthermore, phonemes are subject to context-sensitive rules which govern their phonetic production (i.e., allophones) and constrain the use of phonemes in certain environments. These context-based patterns also highlight the non-linear structure of phonology—namely that phonemes are not only sequenced but also arranged within a hierarchical structure involving multiple tiers of prosodic units (e.g., syllables, morae). Although the specific conceptualization of these suprasegmental units varies across phonological frameworks (Jun, 2005), there is strong support for the psycholinguistic reality of prosodic structure in phonological systems, including allophonic patterns and speech errors by unimpaired speakers and those with developmental and acquired language disorders (Goldsmith, 2011; Macrae & Tyler, 2014; Romani et al., 2011). These and other aspects of language cannot be adequately described by a phonology based exclusively on linear phoneme sequencing.

As our understanding of phonology as a structured language system continues to advance, so must our understanding of the role of phonological structure in typical and atypical acquisition of speech and language. Given the multifaceted complexity of phonological structure that must be acquired during language development, the primary goal of the research presented here is to leverage our understanding of segmental and syllable-level phonology to provide more effective and accessible clinical services for children with language impairments affecting phonological development.

Efficient Intervention for Phonological Disorder

In speech-language pathology, the efficacy and efficiency of assessment and treatment have been focal areas of investigation and are critical for evidence-based service provision (Kamhi, 2006). Treatment for speech sound disorders, which include PD in addition to other functional and organic speech impairments, has perhaps the longest history in the profession. So called “speech doctors” or “speech correctionists” emerged in the late 19th century, providing speech coaching and treatment for individuals with communication challenges related to speech production and/or stuttering. The first “speech correction” organizations were established in the early 20th century with the formation of the National Society for the Study and Correction of Speech Disorders in 1918 and the American Academy of Speech Correction in 1925. The latter would later become the American Speech and Hearing Association in 1947, only formally acknowledging “language” in 1978 when it became the American Speech-Language-Hearing Association (American Speech-Language Hearing Association, 2021; Duchan, 2002).

The history of speech treatment in this profession is relevant to this discussion for a few reasons. First, the relatively longer history of speech (and associated aspects of phonology) may be related to both its salience in communication and its malleability with intervention. Speech sound disorders are noticeable, often immediately upon speaking. This has ramifications for the effect of speech sound disorders on communication as well as psycho-social and emotional development and academic and occupational attainment (Beitchman, Wilson, Brownlie, Walters, Inglis, et al., 1996; Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Felsenfeld et al., 1992, 1994; Lewis et al., 2016; Peterson et al., 2009).

The early impact of speech sound disorders on speech production and intelligibility can be mitigated with treatment, and this has been shown with some degree of consistency across treatment approaches in meta-analyses (Law et al., 2004; Wren et al., 2018). In contrast, outside of expressive vocabulary, the effect of treatment targeting other language domains, such as morphology, syntax, or pragmatics, may not be as consistent (Law et al., 2004). Second, the long-standing history of speech in the profession may impact the novelty of our approaches to assessment and treatment of speech sound disorders. Given a history of effective speech intervention, traditional approaches to speech assessment and especially treatment, dating back to methodologies outlined in Van Riper (1939), are still the most commonly implemented methods of service provision for children with speech sound disorders (McLeod & Baker, 2014).

The immediate functional impact of PD is reduced speech intelligibility (Shriberg & Kwiatkowski, 1994). This alone is a considerable barrier to communication; however, other areas of linguistic and non-linguistic development may also be at risk (e.g., Lewis et al., 2016). Further, the complexity of phonological structure suggests that there is much more at play than misarticulation when children produce words in error or when the acquisition of phonology does not follow the typical trajectory. Consequently, our understanding of phonology and its role in understanding the deficits associated with PD could present novel pathways for improving the efficiency of assessment and treatment by addressing the phonological system, rather than just the surface presentation of speech "errors."

Perspectives on Treatment Efficiency

Approaches to target selection in treatment for PD provide important context for the studies in this dissertation which address treatment efficiency. To provide this background, we proceed with a critical and integrated examination of the literature on approaches to treatment target selection, the aspect of treatment for PD that has most consistently been implicated in impacting treatment efficiency (Kamhi, 2006).

Treatment for PD is provided by a speech-language pathologist or other appropriately qualified clinician with the goal of improving the child's intelligibility by inducing improvement to the child's inaccurately produced structures. Most treatment for PD targets consonant phonemes (e.g., /p/, /k/, /s/, /l/) or consonant clusters (e.g., /kl/, /spl/). Even intervention approaches that target phonological processes (e.g., stopping, final consonant deletion, cluster reduction), natural classes of phonemes (e.g., fricatives, liquids), feature contrasts (e.g., stridency), or phoneme contrasts (e.g., /s/ vs. /θ/) generally target particular phonemes or clusters as exemplars of these broader phonological targets (Baker et al., 2018; Gierut, 1998b; Law et al., 2004; McLeod & Baker, 2014).

Most preschool and early school-age children with PD have multiple production errors affecting many consonants or consonant clusters (e.g., Shriberg & Kwiatkowski, 1994). There is very little research to inform the ideal number of targets to train at a given time, although one comparative study found little difference when multiple or single targets were used (Tyler et al., 1987). The current assumption is that one sound can be trained in isolation or multiple sounds can be trained simultaneously or "cycled" (see Hodson & Paden, 1983). In either case, the intervening clinician must determine which phonological structure or structures to target first, and, subsequently, in what order to target the remaining structures. In the traditional approach (e.g., Rvachew & Nowak, 2001), treatment targets are selected in developmental sequence, based on normative speech acquisition data, so that relatively *simpler* structures are targeted first. Conversely, in a complexity approach (e.g., Gierut & Morrisette, 2012), targets are selected in an inverted developmental sequence so that more *complex* structures are targeted first. As will be discussed, such approaches to target selection may differentially impact the efficiency of treatment for children with PD.

Traditional Approach

A traditional approach to target selection reflects target selection criteria that have been used for decades to treat speech production errors à la Van Riper (1939):

> "In cases where the person makes more than one error, it is well to work with the sounds according to their usual developmental order: first the lip sounds, then the dentals, then the gutturals, then the complicated tongue sounds, and, finally, the blends." (p. 205)

Consequently, a traditional, normative approach to treatment target selection entails first targeting sounds that are early developing according to normative data (Goldman & Fristoe, 2000; McLeod & Crowe, 2018; Smit et al., 1990; Templin, 1957; Wellman et al., 1931) and that the child produces with some error. Because developmentally early-acquired structures are expected to improve quickly (Dyer et al., 1987; Rvachew, 2005; Rvachew & Nowak, 2001; cf., Tyler & Figurski, 1994), this approach facilitates rapid mastery of the target and is expected to improve intelligibility by expediting progression through treatment targets (Rvachew & Nowak, 2001). This approach is also compatible with such well-established developmental theories as Vygotsky's (1978) zones of proximal development because it creates a stepwise pathway upward to the ultimate target of adult-like, intelligible speech production.

In its more recent iterations (e.g., Rvachew & Bernhardt, 2010), a traditional approach pertains not only to the normative developmental sequence of sound acquisition, but also to child-internal factors related to the target, such as a baseline ability to imitate the sound (i.e., stimulability) and level of pre-treatment knowledge (operationalized as percentage of accurate use, contrastive use, or presence in the segmental inventory, among others). In line with the selection of developmentally early-acquired sounds, targets are selected such that child-internal factors would also support faster acquisition of the treated target—namely, stimulability and some pre-treatment knowledge of the sound's use. As with early-acquired phonological structures, sounds for which a child has more pre-treatment knowledge have been shown to improve more rapidly during treatment (Rudolph & Wendt, 2014; Rvachew & Nowak, 2001; Tyler et al., 1993). Thus, targeting simpler (i.e., developmentally earlier acquired and known by the child) phonological structures results in efficient improvement of the targeted structure.

Because this stepwise progression from simpler to more complex phonological structures during treatment mimics the developmental sequence, this approach is expected to be easier to implement, for both the learning child and their clinician. Per Rvachew and Nowak

(2001, p. 610), “it is important that the child not feel frustrated or discouraged by the therapy process.” By targeting earlier-developing sounds for which the child has some pre-existing knowledge or capacity, the target is acquired more quickly and, presumably, with less frustration. As for ease of instruction, few would argue with the claim that it is easier to teach simple subtraction than complex long division; similarly, it should be easier to teach a child to master /k/ than a more complex phoneme, /ɹ/, or complex cluster, /fɹ/. It is noted, however, that there is little experimental support for these claims related to difficulty of instruction. For instance, Rvachew and Nowak (2001) found that children treated for simpler sounds rated therapy just as positively as children treated for more complex sounds.

Rvachew and Bernhardt (2010) appeal to dynamic systems theory (Fogel & Thelen, 1987) to further support a traditional approach to target selection. Dynamic systems theories have existed for some time in the domains of physics, mathematics, biology, and psychology (Thelen & Smith, 2007), and, within the realm of human behavior, dynamic systems have received much attention in research on the acquisition of increasingly complex motor behaviors (e.g., Thelen, 1995). The crux of this theory is the fluid contribution of interacting systems to the emergence of novel, complex behavior. Essentially, dynamic systems theory formalizes “self-organization,” the process by which “[novel] pattern and order emerge from the interactions of the components of a complex system without explicit instructions, either in the organism itself or from the environment” (Thelen & Smith, 2007, p. 259). Crucially, 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 behavioral changes (such as the emergence of a new, more complex sound production pattern). In this scenario, the new behavior is not the product of the one small “trigger” change, but rather the product of accumulated changes prior to and including the trigger.

In a between-subjects experiment with randomized group assignment, Rvachew and Nowak (2001) compared two groups of children with PD (*N* = 48) who received treatment targeting either simpler or more complex targets. They found that children trained with earlier-acquired, most-known targets mastered their targets more rapidly than those trained with later-acquired, least-known targets. They did not find group differences in generalized learning to non-treated phonological structures (i.e., improvement to sounds not targeted during treatment).

Several years later, Rvachew and Bernhardt (2010) re-analyzed data from a subset of six participants in Rvachew and Nowak (2001). They reported that three children treated with simpler targets demonstrated greater improvement to their treated *and* untreated sounds than three children treated with more complex targets (cf. Gierut et al., 1996; Morrisette & Gierut, 2003). Although there are limitations to a retrospective re-interpretation of these data that is incongruent with the inconclusive group outcomes of the original study, the observable improvement in novel, untreated, and even more complex sounds following treatment of a simple target is analogous to the effect of a "trigger" change within a dynamic system. In short, a dynamic systems framework allows for the observation of widespread or cascading improvements in speech-sound production following introduction of a single, relatively simple sound or even strengthening knowledge of a sound already used with some accuracy by the child.

At the core of a dynamic systems framework, however, is the idea that we cannot know or predict which targets are expected to trigger change in the form of novel, more complex phonological productions. Consequently, it seems counter-intuitive that this framework, which eschews the role of input, would be cited in support of any one target selection criterion over another. It could be said that dynamic systems theory does not necessarily support a normative approach over a complexity approach but rather circumvents the role of a particular target entirely and provides a suitable explanation to account for cascading improvements induced by treatment using a target of any type. Nevertheless, the three participants treated with simple targets in Rvachew and Bernhardt (2010) demonstrated improvement in more untreated sounds than the three participants treated with more complex targets (cf. Gierut et al., 1987; Gierut et al., 1996; Powell et al., 1991; Tyler & Figurski, 1994).

Complexity Approach

A complexity approach to target selection is squarely juxtaposed with the target selection criteria of a traditional, normative approach. A complexity approach suggests that the optimal treatment target is linguistically more complex, which means these targets are generally later-developing (Gierut et al., 1996). Similarly, this approach also recommends treatment targets for which the child demonstrates least knowledge and that are produced with little or no accuracy (Gierut et al., 1987), even in imitation (i.e., nonstimulable sounds; Powell et al., 1991).

Opposing target selection criteria in complexity and traditional approaches can be attributed in large part to their diverging views on outcome goals and input. Excepting the relatively more recent suggestion that simple phonological targets can trigger development of untreated, complex sounds (Rvachew & Bernhardt, 2010; cf. Gierut et al., 1996), the original goal of a traditional approach was to facilitate movement through successive treatment targets, with mastery of each target contributing to the child's phonological development (Rvachew & Nowak, 2001; Van Riper, 1939). In contrast, the goal of a complexity approach is not mastery of the target structure, but rather the system-wide impact of improvement to untreated simpler phonological structures (Gierut, 2007).

Language learnability theory, the theoretical framework most often cited in support of a complexity approach, is linguistic rather than behavioral or psychological in its motivation, derived from typological descriptions of human language patterns (Chomsky & Halle, 1968; Clements, 1990; Greenberg, 1978; Ladefoged & Maddieson, 1996; Lindblom & Maddieson, 1988; Maddieson, 1984) and in the language of children across stages of development (Pinker, 1984; Tesar & Smolensky, 1998; Wexler, 1982; Wexler & Culicover, 1980). From these observations, language universals, implicational laws, and markednesss relationships have been deduced, suggesting that language predictably follows expected patterns and restrictions.

The relationship between a complex linguistic structure and its corresponding simpler “prerequisite” structure(s) has been referred to as *markedness* (Prince & Smolensky, 1993). Following this terminology, a marked structure necessarily implies an unmarked structure. A phonological example of this phenomenon would be the relationship between stop and fricative consonants. Typologically, all known languages have stop consonants, but not all languages have fricative consonants. These sound classes maintain a markedness relationship in that no language includes (marked) fricatives that does not also include (unmarked) stops (i.e., fricatives imply stops; Greenberg, 1978; Ladefoged & Maddieson, 1996). Similarly, sonority—a phonological construct related to acoustic intensity—has been associated with a markedness relationship between types of consonant clusters. Adjacent, tautosyllabic consonants whose sonority levels are more distant (e.g., /t/ and /w/ in /tw-/) are considered simple and unmarked relative to those whose sonority levels are close (e.g., /f/ and / ɹ/ in /fɹ-/). Thus, presence of low sonority distance clusters *implies* presence of high sonority distance clusters (Blevins, 1995; Clements, 1990). Because children also maintain lawful linguistic systems (Jakobson, 1968)—and despite their near-constant state of change—acquisition follows these same markedness principles. As such, children mirror cross-linguistic patterns in that they tend to acquire stops before they acquire fricatives (Smit et al., 1990) and high sonority distance clusters before they acquire low sonority distance clusters (Gierut, 1999; McLeod et al., 2001; Smit et al., 1990). Such implicational laws have been documented for a variety of phonological parameters and structures, including those that are often used to identify complex (i.e., marked) treatment targets (shown in Table 1-1; see also Barlow et al., 2011; Gierut, 2007).

Table 1-1. Implicational laws for complex treatment target selection.

Implicational Law			
Marked Structure		**Unmarked Structure**	**Evidence**
Small SD Clusters	⇒	Large SD Clusters	Gierut, 1999
Clusters	⇒	Singletons	Gierut & Champion, 2001
Clusters	⇒	Affricates	Gierut & O'Connor, 2002
Liquids	⇒	Nasals	Dinnsen et al., 1990
Affricates	⇒	Fricatives	Gierut et al., 1994
Fricatives	⇒	Stops	Dinnsen & Elbert, 1984

Note. Adapted from Barlow, Taps, and Storkel (2011) and Gierut (2007). Marked phonological structure (left) implies the corresponding unmarked phonological structure (right). SD = Sonority distance.

In the purest form of learnability theory, these laws can be successfully applied to any adult or child language, accurately predicting the occurrence (or non-occurrence) of phonemes or other phonological or supraphonological structures based on the other elements that exist within that system (for an overview, see Gierut, 2007). In other words, children (both typically developing and with PD) present with lawful linguistic systems subject to the same constraints and implicational relationships that govern the target languages they are acquiring (Dinnsen, 1992; Dinnsen et al., 1990; Jakobson, 1968; Leonard, 1992). The existence of these language universals has spawned a variety of nativist, empiricist, and non-linguistic explanations that are hotly contested and actively evolving (e.g., Cysouw, 2003; Watts & Rose, 2020). Nevertheless, these implicational patterns in adult and child language systems are robustly attested, and a complexity approach ostensibly capitalizes on these patterns to stimulate phonological acquisition.

Crucially—and in contrast to dynamic systems theory—a language learnability framework also emphasizes the role of input in defining an individual's language-specific learning trajectory. For children with PD, this becomes relevant when these language laws are applied to the acquisition of new phonological structures or contrasts between structures in targeted intervention. Per this framework, a given child at a given point in time has a unique and lawful language system that is, generally speaking, a subset of the target adult system. Deliberate exposure to phonological input outside of the bounds of the child's subset system requires the child's language system to expand to accommodate this new input. Consequently, when a child acquires a highly complex structure, they are expected to also acquire simpler prerequisite structures that are required of a lawful phonological system. Thus, language learnability theory supports the prediction that untreated phonological structures can be acquired through deliberate exposure and consequent acquisition (i.e., treatment) of a more marked phonological target. This emphasis on introducing a particular structure as input to the child's developing language system is a considerable divergence from dynamic systems theory, which minimizes the role of a particular input in the induction of change, improvement, or development.

As with the relationship between dynamic systems theory and a normative approach, learnability theory is a broad framework that accommodates the predictions of a complexity approach but is not necessarily exclusive to it. The full range of permissible variations required of the input to allow children to successfully acquire new target forms (Gierut, 2007) is arguably attainable through other treatment approaches. For instance, the cycles approach (Hodson & Paden, 1983), a popular normative approach utilizing systematically cycled targets, exposes the child to a greater breadth of phonological input. By targeting a phonological process rather than one or a few speech sounds, this approach systematically exposes the child to a variety of target speech sounds during each step and, furthermore, does so in a variety of contexts as treatment progresses. In fact, this level of variability has been shown to improve learnability in other language domains (e.g., Aguilar et al., 2018; Plante & Gómez, 2018; Plante et al., 2014), although this effect has not yet been replicated in treatment for phonological disorder (Oglivie, 2019).

Although aspects of learnability can support multiple treatment approaches, a complexity approach is unique in its reliance on implicational laws to improve learning of untargeted structures. Many aspects of complexity have been applied to treatment target selection, summarily categorizable as either child external or child internal. Child-external factors are related to the phonological target's typological, linguistic, or normative characteristics that define its relative complexity. These factors include normative age of acquisition (Gierut et al., 1996) and linguistic markedness (Dinnsen et al., 1990; Dinnsen & Elbert, 1984). These external factors are considered immutable characteristics that would uniformly apply to a given structure. Conversely, child-internal factors are evaluated on a case-by-case basis. These factors pertain to a given child's level of pre-treatment knowledge or ability (*inventory inclusion*: e.g., Gierut & Neumann, 1992; *accuracy*: e.g., Gierut et al., 1987; *stimulability*: e.g., Powell et al., 1991) related to the targeted structure.

The literature base for a complexity approach consists primarily of a series of single-case experimental design studies conducted over a period of 30 years. In most cases, each study was designed to examine one of the aforementioned features of complexity and its subsequent impact on treatment outcomes—specifically, generalized learning of untreated sounds. Prior to these studies, accurate use of the treated sound in non-treatment contexts (i.e., untrained words, sentences, and conversation) or to sounds of the same class (e.g., treatment of stop /k/ improves other phonemes of the same manner class) was considered the extent of possible treatment-induced generalization (e.g., Costello & Onstine, 1976; Elbert & McReynolds, 1975). However, in a complexity framework, treatment of a marked, complex sound is expected to cause simultaneous across-class improvement to other unmarked, simple structures.

For instance, Gierut et al. (1996) conducted two studies examining the efficacy of targeting sounds in developmental and non-developmental sequence. In the second of these studies, the authors found that three preschool-age children with PD who were trained with a later-developing sound (/ɹ/, /θ/, or /s/) demonstrated greater generalized learning of non-treated sounds across manner classes than those trained with an earlier-developing sound (/k/, /g/, or /f/). Other similarly structured studies found that complex targets across a variety of child-internal and child-external target parameters resulted in greater generalized learning of untreated structures. Notably, many of these studies either did not distinguish (Elbert et al., 1984; Gierut, 1998a, 1999) or generally supported the finding that the complex targets themselves may not be acquired as efficiently as simpler targets (Dinnsen & Elbert, 1984; Miccio et al., 1999; Powell et al., 1991)—a finding consistent with a traditional approach.

In summary of this large body of generally small-*N* studies, multiple child-internal and child-external parameters converge to suggest that relatively complex targets are more likely to induce generalized improvement to untreated phonological structures than simpler targets. Thus treatment is expected to lead to simultaneous improvement of both targeted *and* untargeted structures. However, the parameters by which complex targets have been selected vary across studies, which is problematic for their clinical implementation. Notably, there is growing support for the feasibility and efficacy of combining these parameters to select maximally complex targets and induce improvement of untreated structures (Storkel, 2018; Tambyraja & Dunkle, 2014; Taps Richard et al., 2017). However, there is little *comparative* evidence to suggest that, when these complexity parameters are combined, generalization outcomes are better for complex targets than for simple targets. The subsequent task within this framework, then, is to determine which aspects of complexity optimize generalized or system-wide phonological learning. Especially if aspects of complexity carry with them a time-cost related to learning of the treated target, it would be important to determine which parameters can be best combined or which should be prioritized.

Comparative Evidence

Despite their differences, there is compelling evidence to support both a normative approach and a complexity approach to target selection for children with PD. Much of the evidence discussed thus far supports the efficacy of either approach; however, the more interesting and complicated question is which of these approaches or which aspects of these approaches result in better treatment outcomes and enhanced treatment efficiency. We now focus our discussion on the literature that systematically compares aspects of these approaches in order to integrate the most robustly supported features of each.

Rvachew and Nowak (2001) demonstrated that children master the treated target more effectively when the target is simpler (i.e., an earlier-acquired sound with some accurate use prior to treatment), and multiple studies drawing from diverse frameworks support these findings (Dinnsen & Elbert, 1984; Dyer et al., 1987; Powell et al., 1998; Rudolph & Wendt, 2014; Tyler et al., 1993). However, several smaller studies conducted within a language learnability framework found that acquisition of the treated target was not modulated by that target's complexity (Gierut et al., 1987; Gierut et al., 1996; Powell & Elbert, 1984; Powell et al., 1991). Consequently, when the goal is to systematically improve a series of targeted sounds, a normative approach may be most appropriate, although some conflicting evidence leaves this an open question and suggests that other factors likely play a role in the child's ability to master a given sound target trained during treatment. In any case, the most robust finding related to a normative approach to target selection is that sequencing targets from earliest-acquired to latest-acquired allows the child to learn the targets most efficiently and, consequently, the child is likely able to master more treated targets in less time when those targets are simple (e.g., Rvachew & Nowak, 2001). There is currently less support for the claim that treatment of a single simple target results in greater improvement of *untreated* phonological structures than treatment of a single complex target (cf. Rvachew & Bernhardt, 2010).

Gierut and colleagues demonstrated that manipulation of the complexity of the treatment target can predict the type and quantity of generalized learning to untreated sounds. Despite their findings, the implementation of this approach is hindered by the breadth of parameters shown to impact complexity and a lack of experimental evidence to support the optimal combination or prioritization of these parameters. Furthermore, some of these child-internal complexity parameters, such as pre-treatment stimulability or accuracy, may come at the cost of efficiency in treatment target acquisition.

A non-exhaustive survey of the existing literature that *compares* the impact of simple versus complex aspects of child-internal knowledge and child-external linguistic characteristics of the treatment target on a) efficient mastery of the treated structure and b) amount of across-class generalization to untreated phonological structures is displayed in Table 1-2 and Table 1-3, respectively. As discussed above, much evidence has been collected demonstrating the efficacy of a complexity approach, but the comparative evidence for its efficacy above and beyond a normative approach is notably lacking. For instance, direct comparison with a normative approach paradoxically suggests that complex targets are more efficacious (Gierut et al., 1987; Gierut & Morrisette, 2012; Gierut et al., 1996; Powell et al., 1991), that simple targets are more efficacious (Elbert & McReynolds, 1979; Rvachew & Bernhardt, 2010), or that there is no differential improvement to untreated structures between groups of children who learned simple or complex targets (Mota et al., 2007; Pagliarin et al., 2009; Powell & Elbert, 1984; Rvachew & Nowak, 2001). Thus, neither treatment framework, alone, is able to fully account for the patterns of improvement in untreated sounds that have been attested in the growing treatment literature of children with PD.

Table 1-2. Non-exhaustive summary of studies comparing target mastery following treatment of simple versus complex sound structures.

Study	***N***	**Feature or Construct Examined**	**Targ. Mast.**
Dinnsen & Elbert, 1984	4	Knowl: Known targets > Unknown targets	S > C
Forrest, Dinnsen, & Elbert, 1997	10	Knowl: Low substitution variability > High substitution variability	S > C
Forrest, Elbert, & Dinnsen, 2000	14	Knowl: Low substitution variability > High substitution variability	S > C
Gierut & Neumann, 1992	1	Knowl: MinPair w/ 2 absent segments > MinPair w/ 1 absent segment	C > S
Gierut et al., 1996 (Study 1)	3	Targ: Late-acquired > Early-acquired	C > S
Gierut et al., 1996 (Study 2)	6	Targ: Late-acquired = Early-acquired	C = S
Gierut, 1990	3	Targ: MinPair maximal opposition > MinPair minimal opposition	C > S
Gierut, 1991	3	Knowl: MinPair w/ 2 absent segments > MinPair w/ 1 absent segment	C > S
Gierut, 1992	4	Knowl: MinPair w/ 2 absent segments > MinPair w/ 1 absent segment	C > S
Gierut & Dinnsen, 1987	6	Knowl: Most-known targets > Least-known targets	S > C
Gierut, Elbert, & Dinnsen, 1987	6	Knowl: Sequenced targets least-known first > Most-known first	C > S
Kreuger, 2017	6	Knowl: Younger children > Older children	C > S
Powell & Elbert, 1984	6	Targ: Fricative+liquid clusters = Stop+liquid clusters	C = S
Powell et al., 1998	18	Knowl: Phonetic inventory inclusion > Exclusion	S > C
Powell, 1993	6	Knowl: Stim > Non-stim	S > C
Powell, Elbert, & Dinnsen, 1991	6	Knowl: Stim > Non-stim	S > C *
Rudolph & Wendt, 2014	3	Targ: Early-acquired > Late-acquired	S > C
Rvachew & Bernhardt, 2010	6	Targ: Early-acquired > Late-acquired / Knowl: Stim > Non-stim / Least-known > Most-known	S > C
Rvachew & Nowak, 2001	48	Targ: Early-acquired > Late-acquired / Knowl: Stim > Non-stim / Least-known > Most-known	S > C
Tyler et al., 1990	4	Knowl: Presence of covert contrast > Absence of contrast	S > C
Tyler et al., 1993	7	Knowl: Presence of covert contrast > Absence of contrast	S > C

Note. Only participants involved in the named comparison are included in the *N* for each study. Targ = Child-external aspect of the treatment target. Knowl = child-internal knowledge of treatment target. Stim = Stimulable. SD = Sonority distance. MinPair = Minimal pair. S = simpler. C = more complex. Targ. Mast = Target mastery

* Based on this author's re-analysis of data presented in Powell, Elbert, and Dinnsen (1999). Treated sounds that had higher stimulability were closer to mastery by Phase II than those that were non-stimulable or had lower stimulability ($F_{(1,10)} = 7.71$, $p = 0.02$, $R^2 = 0.44$)

Per Table 1-2, there is considerable support for the use of a traditional approach for inducing efficient mastery of the treated target. This pattern was observable for the child-external linguistic complexity of the target and for the child's level of pre-treatment knowledge of the target. Thus, summation of the existing evidence across theoretical frameworks confirms that children are more likely to learn linguistically simple, most-known structures more efficiently than linguistically complex, least-known sounds that are directly targeted in treatment.

Recall that the primary goal of a complexity approach is system-wide improvement, especially to those sounds that were not targeted during treatment. Of those studies which measured across-class generalization to untreated sounds, there is more consistent support for the efficiency of complexity over a traditional approach to induce this type of generalized phonological growth. However, there are important nuances to this evidence. First, it is not consistently shown that the amount of *knowledge* a child demonstrates of the target segment, contrast, or cluster reliably impacts across-class generalization to untreated sounds (Elbert & McReynolds, 1979; Flint & Costello Ingham, 2005; Gierut, 1991; Gierut et al., 1987; Gierut & Neumann, 1992; Powell & Elbert, 1984; Powell et al., 1991; Rvachew & Bernhardt, 2010; Rvachew & Nowak, 2001; Sommers et al., 1967; Williams, 1991). Several studies found greater knowledge of the treated target resulted in more across-class improvement of untreated structures, yet a comparable evidence base also found the opposite—less knowledge of the treated target resulted in more across-class improvement of untreated structures. This includes knowledge in the form of stimulability for production of the target, production accuracy, inclusion in the segmental inventory, and extent of error patterns prior to treatment. Second, there is more consistent evidence that greater *linguistic complexity* of the treatment target (i.e., markedness, normative age of acquisition) may increase the amount of across-class generalization to untreated sounds (Elbert et al., 1984; Gierut, 1990, 1998a, 1999; Gierut & Morrisette, 2012; Gierut et al., 1996; Pagliarin et al., 2009; cf. Rvachew & Bernhardt, 2010). One study (a re-analysis of six participants from a larger study; Rvachew & Bernhardt, 2010) has found greater across-class generalization when the treated target was linguistically simpler, whereas many more studies have replicated the finding that greater across-class generalization when the treated target was linguistically complex.

Table 1-3. Non-exhaustive summary of studies comparing generalization following treatment of simple versus complex sound structures.

Study	*N*	Feature or Construct Compared	Non-Targ.
Dinnsen & Elbert, 1984	4	Targ: Sequenced targets fricatives first = Stops first	C = S *
Elbert & McReynolds, 1979	5	Knowl: Cluster substitution > Cluster reduction	S > C
Elbert & McReynolds, 1985	4	Targ: Sequenced targets fricatives first = Stops first	S = C *
Elbert, Dinnsen, & Powell, 1984	6	Targ: Fricative+liquid clusters > Stop+liquid clusters	C > S
Flint & Costello Ingham, 2005	7	Knowl: Stimulable = Non-stimulable	S = C
Gierut & Morrisette, 2012	10	Targ: Late-acquired sounds in late-acquired words > Earlier-acquired sounds/words	C > S
Gierut & Neumann, 1992	1	Knowl: MinPair w/ 2 absent segments > MinPair w/ 1 absent segment	C > S
Gierut et al., 1996 (Study 2)	6	Targ: Late-acquired > Early-acquired	C = S * / C > S
Gierut, 1990	3	Targ: MinPair maximal opposition > MinPair minimal opposition	C > S
Gierut, 1991	3	Knowl: MinPair w/ 2 absent segments > MinPair w/ 1 absent segment	C > S
Gierut, 1998	4	Targ: Clusters > Singletons	C > S
Gierut, 1999 (Study 1)	6	Targ: Low-SD clusters > High-SD clusters	C > S
Gierut, 1999 (Study 2)	5	Targ: "True" non-/s/-clusters > Adjunct /s/-clusters	C > S
Gierut & Dinnsen, 1987	6	Knowl: Least-known targets > Most-known targets	C > S
Gierut, Elbert, & Dinnsen, 1987	6	Knowl: Sequenced targets least-known first > most-known first	C > S
Mota & Pereira, 2001	2	Targ: MinPair maximal opposition / hierarchically complex features = Modified cycles	S = C
Mota et al., 2007	21	Targ: MinPair maximal opposition = Modified cycles	S = C
Pagliarin et al., 2009	9	Targ: MinPair maximal/multiple opposition > MinPair minimal opposition †	C > S
Powell & Elbert, 1984	6	Targ: Fricative+liquid clusters = Stop+liquid clusters / Knowl: Segment stim. > Non-stim.	C = S / S > C
Powell, Elbert, & Dinnsen, 1991	6	Knowl: Non-stim > Stim	C > S
Rvachew & Bernhardt, 2010	6	Targ: Early-acquired > Late-acquired / Knowl: Stim > Non-stim / Least-known > Most-known	S > C
Rvachew & Nowak, 2001	48	Targ: Early-acquired = Late-acquired / Knowl: Stim = Non-stim / Least-known = Most-known	S = C
Sommers et al., 1967	288	Knowl: Non-stim > Stim ‡	C > S
Williams, 1991	9	Knowl: Cluster substitution > Cluster reduction / Knowl: Accurate segments > Inaccurate	S > C *

Note. Only participants involved in the named comparison are included in the *N* for each study. Non-Targeted Change = Generalized learning to untreated sounds. Targ = Child-external aspect of the treatment target. Knowl = child-internal knowledge of treatment target. Stim = Stimulable. SD = Sonority distance. MinPair = Minimal pair. S = simpler. C = more complex. Non-Targ. = Non-targeted change
* within-class generalization only † moderate-severe and severe participants only

Linguistic complexity may demonstrate a more unidirectional impact than knowledge because of the presumably immutable nature of linguistic complexity, which—barring those less clear complexity or markedness relationships (e.g., Miccio & Ingrisano, 2000)—applies uniformly to the target phoneme, cluster, or contrast across children. Conversely, child-internal knowledge of the target is individualized, difficult to measure consistently, and may be more substantially impacted be other child-internal factors. Because of its variability across participants, the impact of a child's knowledge of the target on across-class generalization may be more difficult to ascertain, especially in the larger, group-design studies. Nonetheless, the evidence for across-class generalization is relatively robust for linguistically complex or later-developing targets, but the same does not reliably hold for targets that are more complex due to the child's level of baseline knowledge.

Moving forward, the optimal constellation of complexity factors to induce change to untreated sounds, developmental factors to promote efficient mastery of the target, and child-internal factors that impact responsiveness to intervention should be explored. Identification of the combination of factors which result in optimal treatment outcomes may be achievable through small-*N*, single-case experimental designs, but these should be conducted with the goal of developing profiles of responsiveness to intervention that can then be escalated upward to ultimately provide the high level of evidence provided by a randomized controlled trial. Furthermore, the available evidence is highly homogeneous in that it almost exclusively pertains to pre-school and early school-age monolingual English-speaking children with moderate to severe PD in the absence of other co-occurring impairment(s). Work in this area must adequately address questions related to treatment target selection in monolingual children who speak languages other than English, in multilingual children, and those of different ages and diverse cognitive-linguistic deficits. Regardless of the population examined, it is crucial that investigation into these questions adequately integrates the available evidence from different—

even competing—frameworks so that we do not continue to rehash contradictory approaches but instead continue to improve target selection strategies for children with PD.

Equitable Intervention for Phonological Disorder

It may be the case that we can strategically target linguistically complex phonological structures to maximize untargeted phonological growth and thus improve the efficiency of treatment for PD. However, if our goal is to improve outcomes for children with PD via efficient service provision, we must also critically examine the scope of children affected by PD and who thus require access to these services. Developmental language impairments, like PD or DLD, affect the mechanisms by which language is acquired (Bishop et al., 2016). Consequently, these impairments affect individuals of all races, socio-economic backgrounds, and who are monolingual or multilingual users of any number of language varieties (Kohnert, 2007; Leonard, 2014). Much of academia has come under increasing scrutiny for a history which has primarily benefitted majority populations in the development and investigation of research questions, and in homogenizing sampling methods (e.g., Bornstein et al., 2013; Miller et al., 2019; Odekunle, 2020). In speech-language pathology, we are increasingly recognizing the equity crisis created by the severely skewed favoring of White, monolingual users of majority language varieties in research that supports evidence-based assessment and treatment for communication disorders (e.g., Miller et al., 2019; Odekunle, 2020).

All young children with a speech or language impairment, including PD, require access to evidence-based, efficient intervention so that they are prepared for better language outcomes and academic success. This requirement is mandated in the United States ("Individuals with Disabilities Education Act," 2004; Turnbull et al., 2009) and endorsed as best practice by the American Speech-Language Hearing Association (2008). Currently, there is an enormous health disparity in access to efficacious treatment of PD between bilingual children who speak Spanish as their first language and their monolingual English-speaking peers. Spanish is spoken in >39 million US households (US Census Bureau, 2018), and consequently many bilingual children are Spanish dominant prior to school age. Emerging bilingual children need access to intervention as soon as impairment or significant delay can be identified (Perry Carson et al., 2003), and this intervention must be provided in their first language when English acquisition has not yet begun (Gutierrez-Clellen, 1999).

The literature resoundingly supports including the first language in speech-language intervention as best practice for supporting a child's bilingual language development. Briefly, PD affects all languages spoken by a child, and optimal treatment addresses their first and second languages (L1 and L2; Yavas & Goldstein, 1998). Prior to English onset, treatment must target the L1 (Gutierrez-Clellen, 1999). Furthermore, strong L1 acquisition (e.g., Spanish) supports L2 acquisition (Cummins, 1991; Gutiérrez-Clellen et al., 2012; Kohnert et al., 2005). When emerging bilinguals are a) treated in Spanish with little guidance as to Spanish-appropriate treatment targets, b) treated in English when this language is not yet spoken by the child, or c) postponed from receiving treatment until school entry or later, these children are systematically excluded from access to the best language and academic outcomes that are accessible to English-speaking monolinguals. This creates a significant health disparity between bilingual and monolingual children.

Also of concern are the silos in which many communication disorders are investigated, as these do not reflect the large caseloads of children with complex constellations of strengths and impairments in communication and related skills. An estimated 34% of children with communication impairments have multiple disorders (Black et al., 2015b). In particular, the traditional relegation of PD to the domain of speech, and exclusionary diagnostic criteria in research related to PD (Bishop et al., 2016; Bishop et al., 2017) may be obscuring its compelling relationship with other impairments, such as DLD and reading disorder (Pennington & Bishop, 2009). For instance, Shriberg and Kwiatkowski (1994) found between 10% and 77%

of 3- to 6-year-old children with speech sound disorders demonstrate co-occurring deficits in other areas of expressive language (i.e., morphology and syntax). Consequently, we cannot ignore the co-occurrence of phonological and morphosyntactic impairments and may better understand the deficits which underlie PD by considering it in conjunction with its associated impairments.

In order to provide clinical services for children with PD that are more efficient and equitable, we must consider how we can improve assessment and treatment approaches as well as the inclusivity of our work as we investigate phonology, phonological development, and clinical methodologies.

Overview of Studies

Studies included in this dissertation leverage segmental and suprasegmental aspects of phonology to advance the efficiency and equity of our approaches to assessment and treatment for children with PD. Efficiency is addressed by considering computational techniques to streamline phonological assessment and by further exploring the role of phonologically complex treatment target selection in broad growth of a child’s phonological system. Equity is addressed via work that considers typical language use and treatment outcomes for children with PD that represent a wider array of cultural, linguistic, and impairment profiles. Each study described herein address one or more aspects of these themes, as outlined in the following.

Chapters 2 and 3 seek to advance the efficiency of more equitable approaches to phonological assessment, which have important implications for the treatment studies described in later chapters. A thorough analysis that includes independent (i.e., not based on comparison to a model or “correct” production) measures allows for a less biased examination of a child’s phonological abilities. In contrast, relational measures require comparison with “correct” productions, a process by which bias towards majoritized language varieties is likely to be introduced. This is blatantly evidenced by the overwhelming amount of norm-referenced standardized assessments with primarily monolingual English-speaking norming populations (Fabiano-Smith, 2019). To facilitate more efficient independent analyses, such as description of phonetic, phonemic, or consonant cluster inventories, Chapter 2 describes a study examining the validity and accuracy of an automated tool to generate independent inventories from transcribed speech samples. Similarly, Chapter 3 describes a study comparing phonemic inventory analysis to consonant accuracy, a more commonly used relational measure.

Chapter 4 addresses both efficiency and equity in assessment, highlighting the problematic clinical division between speech and language via a morphophonological examination of language samples produced by typically developing Spanish-English bilingual children. The goal of this study is to determine whether variable marking of monosegmental tense and agreement morphemes (i.e., third-person singular /-z/ and past tense /-d/) can be predicted by the surrounding phonological environment. Although this is the only study in the dissertation examining a typically developing population, it provides additional evidence for the relationship between phonology and morphology in language sampling—an important assessment and progress monitoring tool for children with speech and language impairments.

Chapter 5 is an initial examination of a complexity-based approach to treatment simultaneously targeting phonology and grammatical morphology for a child with co-occuring PD and DLD. This case study addresses treatment efficiency by examining change in multiple areas of deficit with a single treatment target. Because this study was conducted on-site at an elementary school, and the participant had a more complex impairment spanning multiple linguistic domains, it provides an initial extension of work examining the efficacy of treatment with a complex target to a more accessible setting and an understudied population.

Chapter 6 presents the final study in this dissertation, using a multiple-baseline, single-case experimental design to examine the efficacy of targeting complex phonological structures in Spanish for treatment of PD in Spanish-English bilingual children. This study advances the efficiency of treatment provision by examining differences in broad phonological growth between treatment targeting consonant clusters (relatively more complex) and singleton consonants (relatively less complex). This study may be the first to systematically examine broad treatment outcomes according to Spanish treatment targets of varying complexity, an important step towards offering Spanish-speaking children more equitable access to efficient treatment.

The dissertation concludes with an integrated discussion of the role of phonology in the pursuit of more efficient and equitable service delivery for children with phonologically based language impairments.

References

Aguilar, J. M., Plante, E., & Sandoval, M. (2018). Exemplar variability facilitates retention of word learning by children with specific language impairment. *Language, Speech, and Hearing Services in Schools, 49*, 72-84.

American Speech-Language-Hearing Association. (2008). *Roles and Responsibilities of Speech-Language Pathologists in Early Intervention: Position Statement*. www.asha.org/policy

American Speech-Language Hearing Association. (2021). *History of ASHA*. https://www.asha.org/about/history/

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.

Barlow, J. A., Taps, J., & Storkel, H. L. (2011). Phonological Assessment and Treatment Target (PATT) Selection (Spanish) [Assessment protocol].

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., Inglis, A., & Lancee, W. (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(6), 815-825.

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(6), 804-814.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., Baird, G., Bauer, A., Bellair, J., Boyle, C., Brownlie, E., Carter, G., Clark, B., Clegg, J., Cohen, N., Conti-Ramsden, G., Dockrell, J., Dunn, J., Ebbels, S., . . . Whitehouse, A. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PLoS ONE, 11*(7), 1-26.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Adams, C., Archibald, L., Baird, G., Bauer, A., Bellair, J., Boyle, C., Brownlie, E., Carter, G., Clark, B., Clegg, J., Cohen, N., Conti-Ramsden, G., Dockrell, J., Dunn, J., Ebbels, S., . . . Whitehouse, A. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 58*(10), 1068-1080.

Black, L. I., Vahratian, A., & Hoffman, H. J. (2015a). *Communication disorders and use of intervention services among children aged 3–17 years: United States, 2012*. https://www.cdc.gov/nchs/data/databriefs/db205.pdf

Black, L. I., Vahratian, A., & Hoffman, H. J. (2015b). Communication disorders and use of intervention services among children aged 3–17 years: United States, 2012. *NCHS data brief*(205).

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 206-244). Blackwell.

Bornstein, M. H., Jager, J., & Putnick, D. L. (2013). Sampling in developmental science: Situations, shortcomings, solutions, and standards. *Developmental Review, 33*(4), 357-370.

Brentari, D., Fenlon, J., & Cormier, K. (2018). Sign Language Phonology. In (pp. 691-721). Oxford University Press.

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

Clark, G. M., & Lum, J. A. G. (2017). First-order and higher order sequence learning in specific language impairment. *Neuropsychology, 31*(2), 149-159.

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.

Costello, J., & Onstine, J. (1976). The modification of multiple articulation errors based on distinctive feature theory. *Journal of Speech and Hearing Disorders, 41*, 199-215.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70-89). Cambridge University Press.

Cysouw, M. (2003). Against implicational universals. *Linguistic Typology, 7*(1), 89-101.

Dinnsen, D. A. (1992). Variation in developing and fully developed phonologies. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications* (pp. 191-210). York Press.

Dinnsen, D. A., Chin, S. B., Elbert, M., & Powell, T. W. (1990). Some constraints on functionally disordered phonologies: Phonetic inventories and phonotactics. *Journal of Speech and Hearing Research, 33*(March), 28-37.

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.

Dodd, B., Leahy, J., & Hambly, G. (1989). Phonological disorders in children: Underlying cognitive deficits. *British Journal of Developmental Psychology, 7*(1), 55-71.

Duchan, J. F. (2002). What do you know about your profession's history? *The ASHA Leader, 7*(23), 4-29.

Dyer, K., Santarcangelo, S., & Luce, S. C. (1987). Developmental influences in teaching language forms to individuals with developmental disabilities. *Journal of Speech and Hearing Disorders, 52*(4), 335-335.

Eadie, P., Morgan, A., Ukoumunne, O. C., & Ttofari, K. (2014). Speech sound disorder at 4 years: Prevalence, comorbidities, and predictors in a community cohort of children. *Developmental Medicine and Child Neurology, 57*(6), 578-584.

Elbert, M., & McReynolds, L. V. (1975). Transfer of /r/ across contexts. *Journal of Speech and Hearing Disorders, 40*, 380-387.

Elbert, M., & McReynolds, L. V. (1979). Aspects of phonological acquisition during articulation training. *Journal of Speech and Hearing Disorders, 44*(4), 459-471.

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.

Fabiano-Smith, L. (2019). Standardized tests and the diagnosis of speech sound disorders. *Perspectives of the ASHA Special Interest Groups, 4*(1), 58-66.

Felsenfeld, S., Broen, P. A., & McGue, M. (1992). A 28-year follow-up of adults with a history of moderate phonological disorder: Linguistic and personality results. *Journal of Speech and Hearing Research, 35*, 1114-1125.

Felsenfeld, S., Broen, P. A., & McGue, M. (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: Educational and occupational results. *Journal of Speech and Hearing Research, 37*, 1341-1353.

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.

Fogel, A., & Thelen, E. (1987). Development of early expresive and communicative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology, 23*(6), 747-761.

Gierut, J. A. (1990). Differential learning of phonological oppositions. *Journal of Speech and Hearing Research, 33*(3), 540-549.

Gierut, J. A. (1991). Homonymy in phonological change. *Clinical Linguistics & Phonetics, 5*(2), 119-137.

Gierut, J. A. (1998a). Natural domains of cyclicity in phonological acquisition. *Clinical Linguistics and Phonetics, 12*(6), 481-499.

Gierut, J. A. (1998b). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research, 41*, S85-S100.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research, 42*(3), 708-726.

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16*(1), 6-17.

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.

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.

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.

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.

Gierut, J. A., & Neumann, H. J. (1992). Teaching and learning /θ/: A non-confound. *Clinical Linguistics & Phonetics, 6*(3), 191-200.

Gierut, J. A., & O'Connor, K. M. (2002). Precursors to onset clusters in acquisition. *Journal of Child Language, 29*(3), 495-517.

Gierut, J. A., Simmerman, C. L., & Neumann, H. J. (1994). Phonemic structures of delayed phonological systems. *Journal of Child Language, 21*, 291-316.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2*. American Guidance Service, Inc.

Goldsmith, J. A. (2011). The syllable. In J. A. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The Handbook of Phonological Theory* (pp. 164-196). Blackwell.

Greenberg, J. H. (Ed.). (1978). *Universals of Human Language, Vol. 2: Phonology*. Stanford University Press.

Gutiérrez-Clellen, V., Simón-Cereijido, G., & Sweet, M. (2012, February 1, 2012). Predictors of second language acquisition in Latino children with specific language impairment. *American Journal of Speech-Language Pathology, 21*(1), 64-77.

Gutierrez-Clellen, V. F. (1999). Language choice in intervention with bilingual children. *American Journal of Speech-Language Pathology, 8*(4), 291-302.

Hodson, B. W., & Paden, E. P. (1983). *Targeting Intelligible Speech: A Phonological Approach to Remediation*. College-Hill.

Individuals with Disabilities Education Act. (2004). (20 USC §§ 1431-1432).

Jakobson, R. (1968). *Child Language, Aphasia and Phonological Universals*. Mouton.

Jun, S.-A. (2005). Prosodic typology. In *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 430-458). Oxford University Press.

Kamhi, A. G. (2006). Treatment decisions for children with speech-sound disorders. *Language, Speech, and Hearing Services in Schools, 37*, 271-279.

Kohnert, K. (2007, July 1, 2007). Evidence-based practice and treatment of speech sound disorders in bilingual children. *Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations, 14*(2), 17-20.

Kohnert, K., Yim, D., Nett, K., Fong Kan, P., & Duran, L. (2005). Intervention with linguistically diverse preschool children: A focus on developing home language(s). *Language, Speech, and Hearing Services in Schools, 32*, 153-164.

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.

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*, 924-943.

Leonard, L. B. (1992). Models of phonological development and children with phonological disorders. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, implications*. York Press.

Leonard, L. B. (2014). *Children with Specific Language Impairment* (2nd ed.). MIT Press.

Lewis, B. A., Freebairn, L. A., Hansen, A. J., Stein, C. M., Shriberg, L. D., Iyengar, S. K., & Gerry Taylor, H. (2006). Dimensions of early speech sound disorders: A factor analytic study. *Journal of Communication Disorders, 39*(2), 139-157.

Lewis, B. A., Patton, E., Freebairn, L., Tag, J., Iyengar, S. K., Stein, C. M., & Taylor, H. G. (2016). Psychosocial co-morbidities in adolescents and adults with histories of communication disorders. *Journal of Communication Disorders, 61*, 60-70.

Lindblom, B., & Maddieson, I. (1988). Phonological universals and consonant systems. In L. M. Hyman & C. N. Li (Eds.), *Language, Speech and Mind: Studies in Honour of Victoria A. Fromkin* (pp. 62-78). Routledge.

Macrae, T., & Tyler, A. A. (2014). Speech abilities in preschool children with speech sound disorder with and without co-occurring language impairment. *Language, Speech, and Hearing Services in Schools, 45*(4), 302-302.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge University Press.

McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics, 28*(7-8), 508-531.

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*, 1546-1571.

McLeod, S., van Doorn, J., & Reed, V. A. (2001). Normal acquisition of consonant clusters. *American Journal of Speech-Language Pathology, 10*(2), 99-110.

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.

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*, 214-229.

Miller, A. L., Stern, C., & Neville, H. (2019). Forging diversity-science-informed guidelines for research on race and racism in psychological science. *Journal of Social Issues, 75*(4), 1240-1261.

Morrisette, M. L., & Gierut, J. A. (2003). Unified treatment recommendations: A response to Rvachew and Nowak (2001) [Letter to the Editor]. *Journal of Speech, Language, and Hearing Research, 46*, 382-385.

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.

Odekunle, E. A. (2020). Dismantling systemic racism in science. *Science, 369*(6505), 780.783-781.

Oglivie, T. (2019). The effect of exemplar variability in the treatment of functional speech sound disorders.

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.

Pennington, B. F., & Bishop, D. V. M. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*(1), 283-306.

Perry Carson, C., Klee, T., Carson, D. K., & Hime, L. K. (2003). Profiles of 2-year-olds with delayed language development: Predicting clinical outcomes at age 3. *American Journal of Speech-Language Pathology, 12*(February), 28-39.

Peterson, R. L., Pennington, B. F., Shriberg, L. D., & Boada, R. (2009, October 1, 2009). What influences literacy outcome in children with speech sound disorder? *Journal of Speech, Language, and Hearing Research, 52*(5), 1175-1188.

Pinker, S. (1984). *Language Learnability and Language Development*. Harvard University Press.

Plante, E., & Gómez, R. L. (2018). Learning without trying: The clinical relevance of statistical learning. *Language, Speech, and Hearing Services in Schools, 49*(3S), 710-710.

Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., Lieser, A. M., & Burton, R. (2014). Variability in the language input to children enhances learning in a treatment context. *American Journal of Speech-Language Pathology, 23*(4), 530-545.

Powell, T. W., & Elbert, M. (1984). Generalization following the remediation of early- and later-developing consonant clusters. *Journal of Speech and Hearing Disorders, 49*(2), 211-211.

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.

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.

Prince, A. S., & Smolensky, P. (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers Center for Cognitive Science, Rutgers University.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1239-1257.

Romani, C., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive Psychology, 62*(2), 151-192.

Rudolph, J. M., & Wendt, O. (2014). The efficacy of the cycles approach: A multiple baseline design. *Journal of Communication Disorders, 47*(1), 1-16.

Rvachew, S. (2005). Stimulability and treatment success. *Topics in Language Disorders, 25*(3), 207-219.

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.

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

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders. I: A diagnostic classification system. *Journal of Speech and Hearing Disorders, 47*(3), 226-241.

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

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research, 42*(6), 1461-1481.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa Articulation Norms Project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779-798.

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.

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.

Tambyraja, S.R., & Dunkle, J.T. (2014). Target selection in speech therapy: Is a non-developmental approach more efficient than a developmental approach? *EBP Briefs, 8*(5), 1-9.

Taps Richard, J., Barlow, J. A., & Combiths, P. N. (2017, November). *Applying phonological complexity in the schools: Insights from 32 case studies* American Speech-Language Hearing Convention, Los Angeles, CA.

Templin, M. C. (1957). *Certain language skills in children, their development and interrelationships* (Institute of Child Welfare, Monograph Series 26). University of Minnesota Press.

Tesar, B., & Smolensky, P. (1998). Learnability in Optimality Theory. *Linguistic Inquiry, 29*(2), 229-268.

Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist, 51*(2), 79-95.

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.

Tomblin, J. B., Records, N. L., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research, 40*(6), 1245-1245.

Turnbull, H. R., Huerta, N., Stowe, M., Weldon, L., & Schrandt, S. (2009). *The Individuals with Disabilities Education Act as Amended in 2004*. Pearson Boston, MA.

Tyler, A. A., Edwards, M. L., & Saxman, J. H. (1987). Clinical application of two phonologically-based treatment procedures. *Journal of Speech and Hearing Disorders, 52*, 393-409.

Tyler, A. A., & Figurski, G. R. (1994). Phonetic inventory changes after treating distinctions along an implicational hierarchy. *Clinical Linguistics and Phonetics, 8*(2), 91-107.

Tyler, A. A., Figurski, G. R., & Langsdale, T. (1993). Relationships between acoustically determined knowledge of stop place and voicing contrasts and phonological treatment progress. *Journal of Speech and Hearing Research, 36*, 746-746.

United States Census Bureau. (2018). American Community Survey 5-Year Estimates.

Van Riper, C. (1939). *Speech Correction: Principles and Methods* (1st ed.). Prentice-Hall.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children, 23*(3), 34-41.

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.

Wellman, B., Case, E., Mengert , I., & Bradbury, D. (1931). Speech sounds of young children. *University of Iowa Studies in Child Welfare, 5*, 7-80.

Wexler, K. (1982). A principle theory for language acquisition. In E. Wanner & L. R. Gleitman (Eds.), *Language Acquisition: The State of the Art* (pp. 288-315). Cambridge University Press.

Wexler, K., & Culicover, P. (1980). *Formal Principles of Language Acquisition*. MIT Press.

Williams, A. L. (1991). Generalization patterns associated with training least phonological knowledge. *Journal of Speech and Hearing Research, 34*(4), 722-733.

Wren, Y., Harding, S., Goldbart, J., & Roulstone, S. (2018). A systematic review and classification of interventions for speech-sound disorder in preschool children. *International Journal of Language and Communication Disorders, 53*(3), 446-467.

Wren, Y., Miller, L. L., Peters, T. J., Emond, A., & Roulstone, S. (2016). Prevalence and predictors of persistent speech sound disorder at eight years old: Findings from a population cohort study. *Journal of Speech, Language, and Hearing Research, 59*(4), 647-673.

Yavas, M., & Goldstein, B. (1998). Phonological assessment and treatment of bilingual speakers. *American Journal of Speech-Language Pathology, 7*(2), 49-49.

Chapter 1 includes material from the dissertation author's unpublished integrative paper for the JDP-LCD, Combiths, P. (2018). *Treatment Target Selection for Phonological Intervention*. The dissertation author was the sole author of this paper.

CHAPTER 2:

Automated Phonological Analysis and Treatment Target Selection Using AutoPATT

Automated Phonological Analysis and Treatment Target Selection using AutoPATT

Philip Combiths[a, b]*, R. Amberg[c], G. Hedlund[d], Y. Rose[d], and J. A. Barlow[a]

[a]School of Speech, Language, and Hearing Sciences, San Diego State University, San Diego, USA; [b]Department of Cognitive Science, University of California, San Diego, USA; Alta Loma School District, Rancho Cucamonga, USA; [d]Department of Linguistics, Memorial University of Newfoundland, St. John's, Canada

*Correspondence concerning this article should be addressed to Philip Combiths,

pcombiths@sdsu.edu,

School of Speech, Language, and Hearing Sciences,

San Diego State University, San Diego, CA 92182, USA

Automated analyses of speech samples can offer improved accuracy and timesaving advantages that streamline clinical assessment for children with a suspected speech sound disorder. In this paper, we introduce AutoPATT, an automated tool for clinical analysis of speech samples. This free, open-source tool was developed as a plug-in for Phon (Rose & Hedlund, 2020) and follows the procedures of the Phonological Analysis and Treatment Target Selection protocol (Barlow, Taps, & Storkel, 2010), including extraction of a phonetic inventory, phonemic inventory with corresponding minimal pairs, and initial consonant cluster inventory. AutoPATT also provides suggestions for complex treatment targets using evidence-based guidelines. Automated analyses and target suggestions were compared to manual analyses of 25 speech samples from children with phonological disorder. Results indicate that AutoPATT inventory analyses are more accurate than manual analyses. However, treatment targets generated by AutoPATT should be viewed as suggestions and not used to substitute necessary clinical judgement in the target selection process.

Keywords: phonology; phonological disorder; automated assessment

Introduction

Thorough phonological assessment is critical for identifying the presence, nature, and severity of speech sound disorders (SSDs), and for identifying appropriate treatment targets and goals. However, thorough assessment is time-consuming, and even the recommended 1–1.5 hours of direct assessment can be insufficient (Bleile, 2002; Miccio, 2002; Skahan et al., 2007). This is in addition to time spent post-assessment in analysis, determination of treatment goals, and paperwork, which is frequently reported to be more time-consuming than the assessment itself (Skahan et al., 2007). In total, most SLPs spend between 2 and 2.5 hours in the speech assessment and post-assessment process for children with suspected SSDs (McLeod & Baker, 2014; Skahan et al., 2007). As the global demand for SLPs serving children with SSDs continues to increase (American Speech-Language Hearing Association, 2016, 2020; Jesus et al., 2017; McAllister et al., 2013; Siewert et al., 2014), the efficiency of thorough diagnostic methods for suspected SSDs is an increasingly pressing concern.

In the context of time and resource restraints, short, standardized articulation tests, such as the Goldman-Fristoe Test of Articulation-3 (GFTA-3; Goldman & Fristoe, 2015), are often used to assess speech production, identify SSDs, and even determine treatment targets (Fabiano-Smith, 2019; McLeod & Baker, 2014; Skahan et al., 2007). However, because these types of standardized measures are limited in scope and heavily focused on relational analyses (i.e. comparing the child's productions to a correct target or a normative database), they fail to fully describe a child's speech as it is used in their day-to-day production. These measures do not provide a sample size with sufficient depth or breadth to conduct independent analyses, such as the establishment of phonetic or phonemic inventories (e.g. Barlow & Gierut, 2002; Eisenberg & Hitchcock, 2010; Elbert & Gierut, 1986; Stoel-Gammon, 1985), which describe a child's complete set of speech sound productions or determine their phonemic contrasts

(Combiths et al., 2019; Eisenberg & Hitchcock, 2010; Fabiano-Smith, 2019; Macrae, 2017). Larger elicitation probes or connected speech samples collected through elicitation or spontaneous production are thus increasingly included in speech assessment as more thorough or naturalistic options (Bankson et al., 2017; Bernhardt & Holdgrafer, 2001; Macrae, 2017; Masterson et al., 2005; Miccio, 2002), but time is a concern when clinicians must transcribe and analyse these samples to extract useful diagnostic information from them.

Computerized tools for phonological analysis may offer at least a partial solution to the time investment required for a thorough speech assessment; however, very few SLPs report using these tools (McLeod & Baker, 2014; Skahan et al., 2007). Infrequent adoption of these technologies may be attributable to the limited scope, availability, or accessibility of tools that have been developed. Per Skahan et al. (2007), the most frequently used computerized phonological assessment tool is Hodson's Computerized Analysis of Phonological Patterns (2003). This commercially available software elicits a 50-word speech sample, automates a relational analysis of phonological error patterns (e.g. substitutions, cluster reductions), and provides error-pattern-based treatment target recommendations. Thus, this tool facilitates a rapid analysis of error patterns in English; however, it is not designed for independent analyses or for use with different types of speech samples. A similar program also exists for Malayalam, Computerized Assessment of Phonological Process in Malayalam (Sreedevi et al., 2013). Other commercially distributed tools, such as Logical International Phonetics Program (Oller & Delgado, 2000), Computerized Profiling (Long et al., 2006), and Computerized Articulation and Phonology Evaluation System (Masterson & Bernhardt, 2001) more flexibly facilitate transcription, analysis, and/or treatment target recommendation; however, these programs are no longer maintained and are either unavailable or incompatible with many modern devices.

Otherwise, only a few computerized phonological assessment tools are currently available and compatible with modern hardware. These include Programs to Examine Phonetic and Phonological Evaluation Records (PEPPER; Shriberg, 1990), maintained by the Weissman

Center at the University of Wisconsin-Madison; Ferramentas para Análise Fonológica Automática [Automatic Phonological Analysis Tools] (APAT; Saraiva et al., 2017), maintained by the University of Aveiro, Portugal; and Phon (Rose & Hedlund, 2020; Rose & MacWhinney, 2014), which is part of TalkBank (MacWhinney, 2007) and maintained by Memorial University of Newfoundland. APAT is freely available (at http://acsa.web.ua.pt/) and completes analyses and produces results within Excel, which makes this tool readily accessible to users familiar with that software. Currently, APAT is streamlined for speakers of European Portuguese using samples from the Teste Fonético-Fonológico–Avaliação da Linguagem Pré-Escolar [Phonological Testing–Pre-School Language Assessment] (Mendes et al., 2009) or the Teste de Articulação Verbal [Verbal Articulation Test] (Guimarães et al., 2014). Phon and PEPPER are stand-alone programs with graphical interfaces for transcription and analysis. Both are freely available (at https://www.phon.ca/ and https://phonology.waisman.wisc.edu/, respectively) and can accommodate a variety of speech sample types, including longer samples from independent probes or connected speech, to conduct clinically relevant analyses. Of these programs, Phon has been most recently updated. Because Phon is relatively accessible and actively maintained, its potential for improving the efficiency of speech assessments merits further examination.

To date, Phon has been most frequently used in research (Rose & Stoel-Gammon, 2015); however, it is also appropriate for clinical assessment and monitoring (Byun & Rose, 2016). Through a graphical user interface, Phon allows utterance segmentation (time alignment) as well as orthographic and phonetic transcription of connected speech or elicited samples of any length. Several of these steps can be automated or partially automated within Phon, which includes International Phonetic Alphabet (IPA) dictionaries and syllabification algorithms for multiple languages. These allow automated generation of target (model) transcriptions and phone-by-phone alignments between target and actual forms, all of which are automatically annotated for syllable-level information (e.g. syllable onsets or codas; syllable stress). Phon includes the capacity to conduct acoustic analysis through integration with Praat (Boersma &

Weenink, 2020), and offers a number of clinically useful analyses, especially relational analyses, such as consonant accuracy/Percentage of Consonants Correct (PCC; Shriberg et al., 1997; Shriberg & Kwiatkowski, 1982) and phonological pattern analysis. Additionally, through a scripting language adapted for phonological queries, Phon permits customizable parsing of phonological data. Although certain independent inventory analyses that can provide a more complete description of a child's speech production are not currently integrated into Phon, these can be added with user-created scripts or plug-ins written in JavaScript or Groovy.

As described, Phon can be clinically useful given its ability to accommodate larger speech samples, partial automation of transcription, and built-in relational analyses; however, its current utility could be improved with the capacity to conduct additional independent analyses. Comprehensive independent analyses are often indicated as part of a thorough phonological assessment (e.g. Miccio, 2002; Skahan et al., 2007; Williams, 2015). For example, Phonological Analysis and Treatment Target (PATT) Selection procedures (Barlow et al., 2010) guide clinicians to conduct several independent analyses, including generating a phonetic inventory, an initial cluster inventory, and a phonemic inventory within a generative phonological framework (Chomsky & Halle, 1968). Together, these analyses provide a useful overview of the child's phonological system without relational comparisons to a correct model. From the results of these analyses, PATT procedures provide instructions for identifying gaps in a child's phonological knowledge and selecting relatively complex treatment targets. The recommendation of relatively complex targets is based on research which suggests that treatment targeting complex phonological structures results in greater system-wide phonological growth than targeting simpler structures (Elbert & McReynolds, 1979; Elbert et al., 1984; Flint & Costello Ingham, 2005; Gierut, 1990, 1991, 1998a, 1999; Gierut et al., 1987; Gierut & Morrisette, 2012; Gierut et al., 1996; Gierut & Neumann, 1992; Pagliarin et al., 2009; Powell & Elbert, 1984; Powell et al., 1991; Sommers et al., 1967; Williams, 1991; cf. Rvachew & Bernhardt, 2010).

In order to supplement the clinical utility of Phon and provide more comprehensive independent analyses, we developed AutoPATT (available at https://github.com/rayamberg/AutoPATT) as a Groovy plug-in for Phon. Because PATT steps are procedural in nature (as described below), AutoPATT is able to replicate much of the manual process via automation. Given IPA transcription from a Phon session, AutoPATT automatically generates a phonetic inventory, a set of minimal pairs identifying phonemic contrasts, a phonemic inventory, and an initial cluster inventory. In keeping with PATT protocol, AutoPATT also generates a set of recommended treatment targets based on gaps in a child's phonological knowledge, as identified from the results of its inventory analyses.

Automated procedures for phonological analysis, such as those conducted by AutoPATT and other similar tools, could provide faster and more accurate speech assessment, although this has not been frequently studied. In one existing study, Saraiva et al. (2017) found that computerized APAT results were highly consistent with results derived manually from a standardized phonological assessment. Otherwise, there is a paucity of work in this area. Most automated phonological analyses have not been tested empirically, perhaps because the accuracy of automated procedures is taken for granted. Nevertheless, one cannot assume the accuracy of automated analyses, phonological or otherwise, because computerized processes can and do produce errored results.

Computational error is generally more systematic than human error (Hirschman & Mani, 2003; Strik & Cucchiarini, 2014), which tends to be more sporadic and unpredictable (McBride et al., 2014; Reason, 2000). When unexpected results arise with digital automation, these are usually the result of an error or oversight in the program's specified procedures, as a computer program is literal in its interpretation of instructions. Programs that are tested appropriately can avoid these systematic errors, allowing them to be used repeatedly while yielding results with consistently high levels of dependability. This is something we cannot expect from human operators, especially given the high degree of descriptive precision involved in the computation of many independent analyses, such as those completed by AutoPATT. Similarly, treatment target suggestions could be derived more systematically from automated algorithms, given appropriate and programmatic procedures. In sum, automated processes require testing and validation with realistic datasets to minimize potential systematic error, confirm intended results, and establish accuracy.

The Current Study

The necessity for identifying the accuracy of automated procedures and comparing them against manual procedures motivated the current study. To provide initial validation of AutoPATT analysis results, we compared computerized independent analyses and target selection with AutoPATT to those same procedures completed manually and identified the accuracy of these analyses using 25 speech samples from young children with phonological disorder. With this study, we seek to answer the following questions:

(1) Are automated phonetic, phonemic, and initial cluster inventories, as generated by AutoPATT, comparable to those same analyses conducted manually following PATT procedures?

(2) Does the accuracy of automated phonetic, phonemic, and initial cluster inventories, as generated by AutoPATT, differ from the accuracy of those same analyses conducted manually following PATT procedures?

(3) Are qualitative differences observable between AutoPATT target recommendations and targets generated manually following PATT procedures?

With this work, we contribute to the limited body of research investigating the accuracy of automated phonological analysis. Although treatment target selection is a component of both PATT and AutoPATT, subjective aspects of target selection and differences between manual and automated procedures make their accuracy difficult to quantify. Nevertheless, we observe target selection via both methods and compare them qualitatively.

Method

Participants and Transcriptions

Participants in this study were 25 monolingual English-speaking children (age range = 3;1–6;7; mean age = 4;3) with functional phonological disorder (i.e. impairment in the production, acquisition, or representation of speech sounds with no known cause; Gierut, 1998b) from the Developmental Phonologies Archive of the Learnability Project[1] (Gierut, 2015b). Raw data were narrow phonetic transcriptions of each child's single-word productions from the Phonological Knowledge Probe (PKP; Gierut, 1985), collected prior to their participation in treatment. Reliability for 10% of consonant transcriptions was reported at 93% (Gierut, 2015a). The PKP samples 293 words (for wordlist, see Gierut, 2015c), with a minimum of five opportunities for each English

[1] Archival data were retrieved from the Gierut / Learnability Project collection of the IUScholarWorks repository at https://scholarworks.iu.edu/dspace/handle/2022/20061 The archival data were original to the Learnability Project and supported by grants from the National Institutes of Health to Indiana University (DC00433, RR7031K, DC00076, DC001694; PI: Gierut). The views expressed herein do not represent those of the National Institutes of Health, Indiana University, or the Learnability Project. The author(s) assume(s) sole responsibility for any errors, modifications, misapplications, or misinterpretations that may have been introduced in extraction or use of the archival data.

phoneme, in each permissible word position. The PKP is also designed to elicit minimal pairs from which an individual's phonemic contrasts can be established. To permit analyses of these data with AutoPATT, transcriptions were converted from their archival format to a format compatible with Phon (for further description of this process, see Combiths et al., 2019)

Automated and Manual Data

From these transcriptions, two types of data were derived to compare agreement across manually generated analyses and automated analyses. For the manual analyses, research assistants in a phonology research laboratory were trained to manually complete PATT analysis procedures. Each research assistant demonstrated proficiency with these procedures using a sample dataset prior to contributing to the study. After this training, research assistants completed the PATT for each of 25 samples. PATT assessment procedures include generating:

(1) a phonetic inventory based on a two-time occurrence in the sample, with a corresponding list of English phones missing from the inventory
(2) a list of minimal pairs demonstrating phonemic contrasts
(3) a phonemic inventory derived from a two-time occurrence of minimal pairs, with a corresponding list of English phonemes missing from the inventory
(4) an inventory of word-initial consonant clusters based on a two-time occurrence in the sample, with a corresponding list of English clusters missing from the inventory

Although only these independent inventory analyses were evaluated quantitatively for the purposes of this study, PATT also includes a more involved complexity-based treatment target selection process (see Gierut, 2007; Morrisette et al., 2006; Storkel, 2018). In abbreviated form, this process includes:

(1) determining if any three-element consonant clusters (e.g. /spl-/) are appropriate targets based on their absence in a child's initial cluster inventory and the presence of components of the cluster (e.g. /p/ and /l/) in their phonemic inventories (Gierut & Champion, 2001)
(2) determining if any two-element consonant clusters (e.g. /fɹ-/) are appropriate targets, based on their absence in the initial cluster inventory, and their complexity relative to other English consonant clusters (Gierut, 1999)
(3) in the absence of potential cluster targets, determining a relatively complex singleton target (e.g. /θ/) based on absence from the phonetic inventory (e.g. Gierut et al., 1987), stimulability (e.g. Miccio et al., 1999), frequency, and age of acquisition (e.g. Gierut et al., 1996)

To generate the automated results, the aforementioned analyses and target selection steps were also completed using AutoPATT, which replicates these same procedures. Resultant inventories and sets of suggested targets were arranged such that each segment or cluster in an inventory or set of targets constituted an item for comparison purposes.

To determine accuracy of manual and automated inventories, the "correct" inventories were generated as follows. Each instance of disagreement between manual and automated analyses was reviewed by one of the authors. Referencing PATT procedures and the original raw data, the inclusion of a given segment or cluster in the phonetic, phonemic, cluster, or set of suggested treatment targets was determined. A different author, blind to the initial designations, made accuracy determinations for 20% of the disagreements. When compared, reliability for these designations was 100%. In instances of agreement, the convergence of manual and automated results determined inclusion of that segment or cluster in the corresponding inventory.

Analyses

In order to compare automated analyses to manual analyses, several metrics of interrater reliability were calculated, including percent agreement, Cohen's kappa (Cohen, 1960), and Scott's pi (Scott, 1955). Cohen's kappa and Scott's pi are suitable measures for categorical data from two coders, accounting for the probability of chance agreement in the data (e.g. Mitani & Nelson, 2017). To determine the accuracy of each, these metrics were also calculated between automated and correct inventories and between manual and correct inventories. Mixed effects logistic regression determined the ability of the automated results and manual results to predict correct outcomes, controlling for participant as a random factor.

Results

Agreement

Reliability between AutoPATT and manual analyses are displayed in table 2-1 as percent agreement, Cohen's kappa, and Scott's pi values. Cohen's kappa and Scott's pi values near 0 are indicative of chance agreement (1 indicates perfect agreement). Percent agreement was highest for phonetic inventories (96%), followed by phonemic inventories (89%), and cluster inventories (76%). Despite these results, negative Cohen's kappa and Scott's pi values for all analyses suggest that agreement between AutoPATT and manual analyses are quite poor, given the relatively high probability of chance agreement in these data.

Accuracy

After considering the comparability of AutoPATT and manual analyses, we examined the relationship between results from AutoPATT and the results verified as correct according to PATT protocol, displayed in table 2-2. Here, percent correct was high for all analyses: 100% or nearly 100% for phonetic and phonemic inventories, and 98% for cluster inventories. Cohen's kappa and Scott's pi indicated high agreement with correct results for phonetic, phonemic, and cluster inventories. Logistic regression indicated that, overall, AutoPATT results were a significant predictor of correct results, $z(24, 1033) = 2.75$, $p < 0.01$.

The relationship between results from manual analyses and the results verified as correct was examined in the same fashion, and these results are displayed in table 2-3. Manual analyses were less accurate than AutoPATT analyses, with correct agreement at 96% for phonetic inventories, 89% for phonemic inventories, and 78% for cluster inventories. Cohen's kappa and Scott's pi indicated poor agreement for all analyses. Logistic regression indicated that, overall, manual results were not a significant predictor of correct results, $z(24, 1033) = 0.38$, $p = 0.71$.

Table 2-1. Interrater reliability for AutoPATT and manual analyses as percent agreement, Cohen's kappa, and Scott's pi.

Analysis	*n*	Auto-Manual % Agreement	*SE*	Cohen's Kappa	*SE*	Scott's Pi	*SE*
Phonetic Inventory	552	95.8%	0.009	-0.016	0.207	-0.021	0.209
Phonemic Inventory	398	89.2%	0.016	-0.014	0.146	-0.057	0.152
Cluster Inventory	108	75.9%	0.041	-0.105	0.189	-0.137	0.195
Total	1058	91.3%	0.009	-0.026	0.102	-0.045	0.104

Table 2-2. AutoPATT analysis reliability as percent correct, Cohen's kappa, and Scott's pi.

Analysis	*n*	AutoPATT % Correct	*SE*	Cohen's Kappa	*SE*	Scott's Pi	*SE*
Phonetic Inventory	552	99.8%	0.002	0.908	0.092	0.908	0.092
Phonemic Inventory	398	100.0%	0.000	1.000	0.000	1.000	0.000
Cluster Inventory	108	98.1%	0.013	0.865	0.094	0.865	0.095
Total	1058	99.7%	0.002	0.908	0.053	0.908	0.053

Table 2-3. Manual analysis reliability as percent correct, Cohen's kappa, and Scott's pi.

Analysis	*n*	Manual % Correct	*SE*	Cohen's Kappa	*SE*	Scott's Pi	*SE*
Phonetic Inventory	552	96.0%	0.008	-0.014	0.212	-0.020	0.213
Phonemic Inventory	398	89.2%	0.016	-0.014	0.146	-0.057	0.152
Cluster Inventory	108	77.8%	0.040	0.034	0.174	0.015	0.178
Total	1058	91.6%	0.009	0.017	0.100	0.050	0.087

Qualitative Results

Quantitative analyses captured the overall relationship between AutoPATT and manual analysis results and provided an estimate of the accuracy of each; however, these did not provide insight into the sources of disagreement between manual and automated results or sources of error in either. For this we examined, qualitatively, the nature of discrepancies between AutoPATT and manual analysis results and their errors relative to correct results. These errors are displayed in table 2-4. Furthermore, differences between AutoPATT and manual target selection were only examined qualitatively.

Most disagreements between AutoPATT and manual analyses (approximately 85%) were attributable to omission of a phone, phoneme, cluster, or treatment target from the relevant inventory or set of targets from the manual analysis. For phonetic and cluster inventories, these manual omissions were most common for non-ambient (i.e. not typically occurring in the target language) segments and clusters (e.g. [ʦ], [θw]) or segments and clusters with diacritic markers (e.g. [b̥], [dᵊw]). Manual omissions from the phonemic inventory were frequently related to missing a second occurrence of minimal pairs for a given contrast or not identifying minimal pairs for non-ambient segments. Most of these omissions were classified as errors in the manual analysis.

Instances in which AutoPATT omitted an inventory item that was included in manual results were less common. These were most often attributable to differences in the interpretation of a phone or cluster for the purposes of inventory inclusion. For instance, [kj] occurred two or more times as a word-medial cluster in several participants' productions. On three occasions, research assistants included [kj] in the initial cluster inventory, based on these word-medial occurrences. However, PATT guidelines specify that inclusion in the cluster inventory should be based on a two-time occurrence of the cluster in word-initial position, making inclusion of [kj] in those cases an error. In a different example, [dð] occurred multiple times in the data. Based on its patterning in the samples, this was most likely meant to represent a dentalized affricate (i.e. [d͡ð]). Research assistants correctly interpreted these transcriptions as affricates, whereas the AutoPATT analysis erroneously interpreted these transcriptions as occurrences of clusters.

Table 2-4. Unique inventory errors.

Phonetic Inventory		*Phonemic Inventory*		*Cluster Inventory*	
Error Source	Errored Item	Error Source	Errored Item	Error Source	Errored Item
manual omission	ʣ	manual addition	ɾ	manual addition	ʔj
manual omission	ʤ	manual omission	wʴ	manual omission	dᵊw
manual omission	nː	manual omission	ʤ	manual omission	bᵊw
manual addition	ʦ̪	manual omission	b̥	manual omission	fw
manual omission	d̪	manual omission	p	manual omission	sw
manual omission	ð̥	manual omission	z	manual omission	bw
manual omission	ʣ̪	manual omission	t̚	manual omission	dwʴ
manual omission	l	manual omission	m	manual addition	kj
manual omission	f	manual omission	h	manual omission	θw
manual omission	jᵊ	manual omission	ɹ	manual omission	gw
manual addition	ʣ̪	manual omission	k	manual omission	dw
manual omission	dᵊ	manual omission	w	manual addition	tw
manual addition	b̥	manual omission	g	manual addition	dw
manual addition	ð	manual omission	b	manual addition	θn
manual addition	ʧ	manual omission	ʦ̪	manual omission	ʃn
manual omission	d̥	manual addition	b̥	AutoPATT addition	dð
manual omission	nᵗ	manual omission	d̥	AutoPATT addition	tθ
manual omission	ɾ	manual omission	z̥		
manual omission	ɹᵗ	manual omission	ʃ		
manual omission	t̚	manual omission	ʔ		
manual omission	ʔ	manual omission	ʦ		
AutoPATT omission	dð	manual addition	v		
		manual omission	s̪		
		manual omission	s		
		manual omission	ʣ̪		
		manual omission	ʧ		
		manual omission	ɾ		
		manual omission	j		
		manual omission	ʣ		
		manual omission	bʴ		
		manual omission	t̚		
		manual omission	tʰ		

Note: Only unique errors are displayed. Repeated errors are only listed once.

The number of manually selected treatment targets differed greatly across manual and automated procedures. This is primarily because most research assistants indicated only one treatment target, whereas AutoPATT provided a list of targets when multiple targets were appropriate. However, the manually selected target was always included in the set of potential targets identified by AutoPATT, with one exception. For one sample, the manually selected treatment target was /skw-/. However, AutoPATT missed this three-element cluster target because it erroneously considered an occurrence of a two-element cluster with a diacritic, [s̪t-], as an instance of a three-element cluster, eliminating potential three-element cluster targets. Identification of these discrepancies and errors provided useful information for future revisions of the AutoPATT algorithm, as discussed in the next section.

Discussion

In this study, we compared the results of manual and automated inventory analysis and treatment target selection, following PATT procedures (Barlow et al., 2010). We first discuss the quantitative and qualitative results for generation of inventories separately from the qualitative results for treatment target selection, as our analyses and findings differed in these areas, and they diverge in their relevant considerations.

Results indicate that automated generation of phonetic, phonemic, and cluster inventories using AutoPATT is not equivalent to these same inventories generated manually by undergraduate and graduate students with training in these procedures, at least when using narrowly transcribed speech samples from young children with phonological disorder. Low percent agreement between manual and automated inventory analyses were confirmed by near-zero Scott’s pi and Cohen’s kappa values. However, accuracy and qualitative error analyses revealed that disagreements were primarily attributable to human error in the manually generated analyses, most frequently omission of a phone, phoneme, or cluster that was accurately identified by AutoPATT. Specifically, AutoPATT-generated inventories were 98–100% accurate, whereas manually generated inventories were 78–96% accurate. Thus, AutoPATT may be a more accurate and consistent means of generating inventories for speech analysis than manual procedures.

In addition to generating inventory analyses, PATT and AutoPATT both include procedures for selecting complex treatment targets (e.g. Morrisette et al., 2006), based primarily on an individual's phonetic, phonemic, and initial cluster inventories. Research assistants frequently indicated only a single suggested target, and AutoPATT indicated a set of potential targets where applicable. Although automated procedures may be able to provide a short list of potentially appropriate targets, AutoPATT is ultimately unable to incorporate the myriad factors involved in determining a single target, based on independent inventory analyses alone. Furthermore, we identified a systematic error in the AutoPATT algorithm that led to misinterpretation of a two-element cluster with a diacritic [s̪t-] as a three-element cluster for the purposes of target selection—although this was addressed in subsequent revisions to the program, as described in the next section. From these observations, we conclude that AutoPATT may be able to streamline complex target selection by narrowing the pool of potential targets, but it is not a substitute for necessary clinical judgement in treatment target selection, and its suggestions must be reviewed against assessment results in case of unexpected errors.

Clinical Implications

Automated phonological analyses, such as the generation of phonetic, phonemic, and cluster inventories, show promise as accurate means of describing an individual's phonological system. As shown here, automated inventories can be generated with less error than those created manually. Human error is a well-documented phenomenon in phonological analysis for research purposes (e.g. Shriberg & Lof, 1991); however, it is less frequently addressed in the clinical domain. For phonological analysis, relatively high proficiency with IPA notation is required, in addition to some knowledge of phonological theory. Even when clinicians are able to conduct these analyses, they are unlikely to have access to the time and resources available to research assistants in a phonology laboratory which permit them to work carefully and review their analyses for errors. Assuming accurate digital transcriptions, automated analyses also allow effective archival of an individual's speech production abilities, which can be referenced and analysed repeatedly. Because analysis procedures are applied identically across datasets, results can be compared over time with confidence that the analyses were conducted consistently. Consequently, clinicians might stand to benefit considerably from the greater accuracy and consistency of automated phonological analyses for speech assessment.

These automated analyses could be more efficient than manually completed analyses, although comparison of the time spent preparing and conducting manual and automated analyses still requires direct investigation. AutoPATT analyses require a speech sample to be transcribed in Phon. Transcriptions may be completed in Phon from a video or audio recording or by hand. If transcription is completed by hand, it requires data entry, either directly into Phon using its built-in IPA map or indirectly with another IPA typing tool. For a clinician or researcher experienced with Phon and digital IPA transcription, this transfer can be completed in 30 minutes for a sufficiently thorough sample, such as the 293-word samples in this study, but it could take an hour or more in other circumstances. Although digital transcription requires an initial time investment, phonological analysis software may offer a significant return on investment, as any number of relational and independent analyses can be completed and repeated in minutes. For comparison, trained students and research assistants typically spent 1–2 hours to complete manual PATT assessment and target selection procedures.

Although a pool of treatment target options, as generated by AutoPATT, can be a useful tool for clinicians seeking to identify relatively complex treatment targets for a child with phonological disorder, these suggestions are based on limited, one-dimensional inventory analyses. The onus of treatment target selection still lies on the clinician who may choose to consider these or other target options in the context of a complete assessment, which may include automated analysis of a speech sample, but should also include a variety of other assessment measures (Fabiano-Smith, 2019; Kamhi, 1992; McLeod & Baker, 2014; Miccio, 2002).

Limitations and Future Directions

Although AutoPATT was shown to be relatively accurate for inventory analyses in its current state, there remain areas for improvement. Research assistants following manual PATT procedures at times interpreted initial consonant clusters as exclusively word-initial and other times included syllable-initial (word-medial) clusters in that category. AutoPATT considered only word-initial clusters for the initial cluster inventory. Indeed, syllable-initial clusters may be appropriate to include in the initial cluster inventory, and both PATT and AutoPATT procedures could be updated accordingly. Also, at least one instance of systematic error in the AutoPATT target selection algorithm was identified, although this only impacted target selection for one participant, and the error has since been corrected.

Currently, AutoPATT does not consider an individual's stimulability for sounds absent from the phonetic inventory in the target selection process because this cannot be determined from a single-word sample. However, stimulability is a relevant consideration in target selection (e.g. Miccio et al., 1999). Similarly, substantial discrepancies between manual and automated treatment target selection highlight the critical role of clinical judgement in this process. Since completion of this study, AutoPATT has been revised to address the error in target selection described above, to display more detailed information explaining the characteristics of the sampled phonological system that resulted in the given set of suggested complex targets, and to clarify the utility of stimulability testing and other analysis tools in conjunction with clinical judgement for evidence-based assessment and treatment target selection.

Replication of this work would improve our understanding of the utility of these automated analysis tools. Extension of this work to other computerized assessment tools, to other populations, and with more heterogeneous samples would be especially beneficial, as the current findings are only generalizable to AutoPATT analyses of samples from young monolingual English-speaking children with phonological disorder. Furthermore, PATT procedures exist for Spanish, and AutoPATT was developed for both English and Spanish;

however, the accuracy of AutoPATT analyses in Spanish remains to be examined. As these automated analyses are revised, they will require additional validation. This iterative process will continue to improve the available repertoire of clinically appropriate tools for phonological analysis.

Conclusion

Automated speech analysis is an emerging area of academic and clinical interest, and it is increasingly considered a useful tool for clinical speech assessment. Clinicians may choose to expand their repertoire of phonological assessment tools to include automated analyses; however, the accuracy and validity of these tools should be taken into consideration, and automated results should be interpreted in the context of other conventional assessment tools and each client’s unique set of personal circumstances and priorities.

Acknowledgements

We thank the members of the Phonological Typologies Lab at San Diego State University, especially Sarah Covington, Stephanie Nguyen, Brooke Hendricks, Taylor Kubo, Diana Carreon, and Teresa Roquet for their assistance with this project.

Disclosure Statement

The fifth author is an author of the Phonological Analysis and Treatment Target (PATT) Selection protocol. The first, second, and fifth authors are authors of AutoPATT. The third and fourth authors are authors of Phon. These tools and software programs are distributed freely. The authors have no relevant financial disclosures.

Funding

This work was supported by the National Institutes of Health under Grants F31 DC017697, R21 DC01720, and R01 HD051698-11.

References

American Speech-Language Hearing Association. (2016). *2016 Schools Survey report: SLP caseload characteristics*. http://www.asha.org/research/memberdata/schoolssurvey/

American Speech-Language Hearing Association. (2020). *Supply and demand resource list for speech-language pathologists*. https://www.asha.org/uploadedFiles/Supply-Demand-SLP.pdf

Bankson, N.W., Bernthal, J.E., & Peter Flipsen, J. (2017). Assessment: Data collection. In J. E. Bernthal, N. W. Bankson, & J. Peter Flipsen (Eds.), *Articulation and Phonological Disorders: Speech Sound Disorders in Children* (pp.150-175). Pearson.

Barlow, J.A., & Gierut, J.A. (2002). Minimal pair approaches to phonological remediation. *Seminars in Speech and Language, 23*, 57-68. https://doi.org/10.1055/s-2002-24969

Barlow, J.A., Taps, J., & Storkel, H.L. (2010). Phonological Assessment and Treatment Target (PATT) Selection [Assessment protocol].

Bernhardt, B.H., & Holdgrafer, G. (2001). Beyond the Basics I. *Language, Speech, and Hearing Services in Schools, 32*(1), 18-27. https://doi.org/doi:10.1044/0161-1461(2001/002)

Bleile, K.M. (2002). Evaluating articulation and phonological disorders when the clock is running. *American Journal of Speech-Language Pathology, 11*(3), 243-249. https://doi.org/10.1044/1058-0360(2002/026)

Boersma, P., & Weenink, D. (2020). *Praat: Doing Phonetics by Computer.* (Version 6.1) [Computer program].

Byun, T.M.A., & Rose, Y. (2016). Analyzing clinical phonological data using Phon. *Seminars in Speech and Language, 37*(2), 85-105. https://doi.org/10.1055/s-0036-1580741

Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Combiths, P.N., Barlow, J.A., & Sanchez, E. (2019). Quantifying phonological knowledge in children with phonological disorder. *Clinical Linguistics & Phonetics*, 1-14. https://doi.org/10.1080/02699206.2019.1584247

Eisenberg, S.L., & Hitchcock, E.R. (2010). Using standardized tests to inventory consonant and vowel production: A comparison of 11 tests of articulation and phonology. *Language, Speech, and Hearing Services in Schools, 41*, 488-503. https://doi.org/10.1044/0161-1461(2009/08-0125)

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

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*, 309-317. https://doi.org/10.1044/jshd.4903.309

Elbert, M.F., & Gierut, J.A. (1986). *Handbook of clinical phonology: Approaches to assessment and treatment.* Singular Group.

Fabiano-Smith, L. (2019). Standardized tests and the diagnosis of speech sound disorders. *Perspectives of the ASHA Special Interest Groups, 4*(1), 58-66. https://doi.org/10.1044/2018_pers-sig1-2018-0018

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*, 53-63. https://doi.org/10.1044/cicsd_32_S_53

Gierut, J.A. (1985). *On the relationship between phonological knowledge and generalization learning in misarticulating children* [Doctoral dissertation, Indiana University]. Bloomington, IN.

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 and 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*, S85-S100.

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*, 6-17. https://doi.org/10.1044/1058-0360(2007/XXX)

Gierut, J.A. (2015a). Experimental designs and protocols. In *Learnability Project Working Papers*. Indiana University. https://doi.org/10.5967/K8CC0XMD

Gierut, J.A. (2015b). *Learnability Project*. Indiana University. https://scholarworks.iu.edu/dspace/handle/2022/20061

Gierut, J.A. (2015c). Phonological protocols. In *Learnability Project Working Papers* (pp.1-50). Indiana University. https://doi.org/10.5967/K87P8W9V

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/doi:10.1017/S0142716411000294

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/0161-1461/96/2703-0215

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

Goldman, R., & Fristoe, M. (2015). *Goldman-Fristoe Test of Articulation 3*. Pearson Clinical Assessment.

Guimarães, I., Birrento, C., Figueiredo, C., & Flores, C. (2014). *TAV–Teste de Articulação Verbal [Test of Verbal Articulation]* Oficina Didáctica.

Hirschman, L., & Mani, I. (2003). Evaluation. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics* (pp.414-429). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199573691.013.014

Jesus, L.M.T., Martinez, J., Valente, A.R., & Costa, M.C. (2017). Speech and language therapy service delivery: Overcoming limited provision for children. *Public Health, 151*, 39-50. https://doi.org/10.1016/j.puhe.2017.06.013

Kamhi, A.G. (1992). Phonological assessment and treatment: The need for a broad-based model of phonological disorders. *Language, Speech, and Hearing Services in Schools, 23*, 261-268. https://doi.org/10.1044/0161-1461.2303.261

Long, S., Fey, M., & Channell, R. (2006). *Computerized Profiling.* (Version 9.7. 0) [Computer software] Case Western Reserve University.

Macrae, T. (2017). Stimulus characteristics of single-word tests of children's speech sound production. *Language, Speech, and Hearing Services in Schools, 48*(4), 219-233. https://doi.org/10.1044/2017_LSHSS-16-0050

MacWhinney, B. (2007). The TalkBank project. In J. C. Beal, K. P. Corrigan, & H. L. Moisl (Eds.), *Creating and Digitizing Language Corpora* (pp.163-180). Springer.

Masterson, J., & Bernhardt, B. (2001). *Computerized Articulation and Phonology Evaluation System (CAPES).* (Version 4) [Computer program] The Psychological Corporation.

Masterson, J.J., Bernhardt, B.H., & Hofheinz, M.K. (2005, August 1, 2005). A comparison of single words and conversational speech in phonological evaluation. *American Journal of Speech-Language Pathology, 14*(3), 229-241. https://doi.org/10.1044/1058-0360(2005/023)

McAllister, L., Wylie, K., Davidson, B., & Marshall, J. (2013). The World Report on Disability: An impetus to reconceptualize services for people with communication disability. *International journal of speech-language pathology, 15*(1), 118-126. https://doi.org/10.3109/17549507.2012.757804

McBride, S.E., Rogers, W.A., & Fisk, A.D. (2014, 2014/11/02/). Understanding human management of automation errors. *Theoretical Issues in Ergonomics Science, 15*(6), 545-577. https://doi.org/10.1080/1463922X.2013.817625

McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics, 28*(7-8), 508-531. https://doi.org/10.3109/02699206.2014.926994

Mendes, A., Afonso, E., Lousada, M., & Andrade, F. (2009). *Teste Fonetico–Fonológico-Avaliação da Linguagem Pré-Escolar (TFF-ALPE) [Phonological Testing–Pre-School Language Assessment]*. Designeed.

Miccio, A.W. (2002). Clinical problem solving: Assessment of phonological disorders. *American Journal of Speech-Language Pathology, 11*, 221-229.

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

Mitani, A.A., & Nelson, K.P. (2017). Modeling agreement between binary classifications of multiple raters in R and SAS. *Journal of Modern Applied Statistical Methods, 16*(2), 277-309. https://doi.org/10.22237/jmasm/1509495300

Morrisette, M.L., Farris, A.W., & Gierut, J.A. (2006). Applications of learnability theory to clinical phonology. *Advances in Speech-Language Pathology, 8*, 207-219. https://doi.org/10.1080/14417040600823284

Oller, K., & Delgado, R. (2000). *Logical International Phonetics Program* (Version 2) [Computer software] Intelligent Hearing Systems.

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

Powell, T.W., & Elbert, M. (1984). Generalization following the remediation of early- and later-developing consonant clusters. *Journal of Speech and Hearing Disorders, 49*(2), 211-211. https://doi.org/10.1044/jshd.4902.218

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*, 1318-1328. https://doi.org/10.1044/jshr.3406.1318

Reason, J. (2000). Human error. *Western Journal of Medicine, 172*(6), 393-396. https://doi.org/10.1136/ewjm.172.6.393

Rose, Y., & Hedlund, G. (2020). *Phon.* (Version 3.1) https://www.phon.ca/

Rose, Y., & MacWhinney, B. (2014). The PhonBank project: Data and software-assisted methods for the study of phonology and phonological development. In J. Durand, U. Gut, & G. Kristoffersen (Eds.), *The Oxford Handbook of Corpus Phonology* (pp.380-401). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199571932.013.023

Rose, Y., & Stoel-Gammon, C. (2015). Using Phonbank and Phon in Studies of Phonological Development and Disorders. *Clinical Linguistics & Phonetics, 29*(8-10), 686-700. https://doi.org/10.3109/02699206.2015.1041609

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

Saraiva, D., Lousada, M., Hall, A., & Jesus, L.M.T. (2017). Paediatric Automatic Phonological Analysis Tools (APAT). *Logopedics Phoniatrics Vocology, 42*(4), 153-159. https://doi.org/10.1080/14015439.2016.1237544

Scott, W.A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly, 19*(3), 321-325. https://doi.org/10.1086/266577

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*, 708-722. https://doi.org/10.1044/jslhr.4004.708

Shriberg, L.D. (1990). *PEPPER: Programs to Examine Phonetic and Phonological Evaluation Records*. [Computer program] Lawrence Erlbaum Associates.

Shriberg, L.D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders, 47*, 256-270. https://doi.org/10.1044/jshd.4703.256

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

Siewert, R., Boyce, S., Seselja, Z., Bilyk, C., Brown, C., Peris, N., Reynolds, L., Smith, D., Thorp, H.L., Urquhart, A., & Moore, C. (2014). The demand for speech pathology services in Australia. In *Prevalence of Different Types of Speech, Language and Communication Disorders and Speech Pathology Services in Australia [Senate report]* (pp.33-43). Commonwealth of Australia.

Skahan, S.M., Watson, M., & Lof, G.L. (2007). Speech-language pathologists' assessment practices for children with suspected speech sound disorders: Results of a national survey. *American Journal of Speech-Language Pathology, 16*(3), 246-259. https://doi.org/10.1044/1058-0360(2007/029)

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*, 428-437. https://doi.org/10.1044/jshr.1003.428

Sreedevi, N., John, M., & Chandran, S. (2013, 2013/08//). Computerized Assessment of Phonological Processes in Malayalam (CAPP-M).

Stoel-Gammon, C. (1985). Phonetic inventories, 15-24 months: A longitudinal study. *Journal of Speech and Hearing Research, 28*, 505-512. https://doi.org/10.1044/jshr.2804.505

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

Strik, H., & Cucchiarini, C. (2014). On automatic phonological transcription of speech corpora. In J. Durand, U. Gut, & G. Kristoffersen (Eds.), *The Oxford Handbook of Corpus Phonology* (pp.89-109). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199571932.013.001

Williams, A.L. (1991). Generalization patterns associated with training least phonological knowledge. *Journal of Speech and Hearing Research, 34*, 722-733. https://doi.org/10.1044/jshr.3404.733

Williams, A.L. (2015). Assessment and intervention from a systemic perspective. In C. Bowen (Ed.), *Children's Speech Sound Disorders* (2nd ed., pp.199-203). Wiley-Blackwell.

Chapter 2, in full, is a reprint of the manuscript, 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.* The dissertation author was the primary investigator and author of this paper.

CHAPTER 3:

Quantifying Phonological Knowledge in Children with Phonological Disorder

CLINICAL LINGUISTICS & PHONETICS
2019, VOL. 33, NOS. 10–11, 885–898
https://doi.org/10.1080/02699206.2019.1584247

Taylor & Francis
Taylor & Francis Group

Check for updates

Quantifying phonological knowledge in children with phonological disorder

Philip N. Combiths [a], Jessica A. Barlow[b], and Emilie Sanchez[c]

[a]Language and Communicative Disorders, San Diego State University, University of California, San Diego, CA, USA; [b]Speech, Language, and Hearing Sciences, San Diego State University, CA, USA; [c]San Diego Unified School District, San Diego, CA, USA

ABSTRACT

Generative phonologists use contrastive minimal pairs to determine functional phonological units in a language. This technique has been extended for clinical purposes to derive phonemic inventories for children with phonological disorder, providing a qualitative analysis of a given child's phonological system that is useful for assessment, treatment, and progress monitoring. In this study, we examine the single-word productions of 275 children with phonological disorder from the Learnability Project (Gierut, 2015b) to confirm the relationship between phonemic inventory – a measure of phonological knowledge – and consonant accuracy – a quantitative, relational measure that directly compares a child's phonological productions to the target (i.e. adult-like) form. Further, we identify potential percentage accuracy cutoff scores that reliably classify sounds as in or out of a child's phonemic inventory in speech-sound probes of varying length. Our findings indicate that the phonemic function of up to 90% of English consonants can be identified from percentage accuracy for preschool-age children with phonological disorder when a sufficiently large and thorough speech sample is used.

ARTICLE HISTORY
Received 24 January 2019
Revised 11 February 2019
Accepted 14 February 2019

KEYWORDS
Phonology; phonological disorder; assessment; phonemic inventory; accuracy

Introduction

Phonological disorder (PD) is one of the most prevalent communication disorders in young children, with prevalence estimates in the range of 7–11% for children at five years of age (Law, Boyle, Harris, Harkness, & Nye, 2000). This developmental impairment of unknown aetiology occurs independently of another primary motivating condition and functionally impairs the development, manipulation, and production of the phonological units of language. Consequently, this form of developmental communication disorder prevents acquisition of phonological skills in a timely manner, which can impact communication and literacy skills and later academic, socio-emotional, and occupational outcomes (Beitchman et al., 1996; Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Felsenfeld, Broen, & McGue, 1992, 1994; Lewis et al., 2016; Peterson, Pennington, Shriberg, & Boada, 2009). Given the prevalence and impact of this impairment, researchers and clinicians alike are continually exploring techniques for assessment and progress

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

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/iclp.

monitoring to better capture the phonological abilities of children with PD. We begin our discussion of this topic by describing and comparing two such measures below.

Phonemic inventory analysis

Speech-language pathologists (SLPs) and researchers utilize a variety of measures to assess productive knowledge of speech sounds in children with PD. One such measure used to qualitatively describe a child's functional phonological knowledge is the phonemic inventory. This unique, linguistically motivated measure ostensibly captures a child's functional use of contrastive speech sounds (i.e. phonemes). The methodology for deriving a phonemic inventory originated in a generative linguistics framework and has been used to provide phonological descriptions of fully formed adult languages (e.g. Voegelin, 1957). In order to characterize those phonemes that are used contrastively to distinguish words among speakers of a given adult language, phonologists require semantically distinct word pairs with minimal phonological contrast (i.e. minimal pairs) to demonstrate a speech sound's phonemic function (e.g. Chomsky & Halle, 1968). For instance, the words 'car' /kɑɹ/ and 'tar' /tɑɹ/ form a minimal pair in English because they differ by only a single segment, demonstrating that the contrast between /k/ and /t/ is sufficient to distinguish words. Thus, /k/ and /t/ are contrastive phonemes in English.

In discussion of phonemes and phonemic inventories within a generative phonology framework, it is important to emphasize that a phoneme is an abstraction that describes a categorical representation of a functional, contrastive unit within a word. These abstract units can have a number of phonetic expressions, all recognized as variants of the same categorical phoneme by speakers of the same language. For instance, the categorical phoneme /p/, in English, can be produced as [p˭] (e.g. [sp˭un] 'spoon'), [pʰ] (e.g., [pʰɑt] 'pot'), or [p̚] (e.g. [stɑp̚] 'stop') – any of these productions would be understood as productions of the phoneme /p/ by a native speaker of English. Furthermore, although the presence of minimal pairs is the primary evidence for establishing phonemic status, it is only one part (albeit an important one) of a larger process involving a comprehensive description of the phonetic environments in which a given sound occurs and its phonetic similarity (or dissimilarity) to other sounds of that language.

The phonemic inventory and its corresponding methodology have subsequently been extended for clinical purposes to derive the phonemic inventories of individual children with PD, although the criterion for phonemic status has been simplified such that the primary requirement is the presence of two minimal pairs (i.e. four words in total) to establish a speech segment as a phoneme in a given child's phonemic inventory (Barlow & Gierut, 2002; Dinnsen, 1984; Gierut, Simmerman, & Neumann, 1994). Because children in the process of language development demonstrate unique and dynamic phonological systems (Fry, 1967; Jakobson, 1968), their phonemic inventories are likewise varied and subject to change over time. Importantly, the phonemic inventory is also considered an independent measure because it neutrally describes the functional phonemes a child uses, without reference to their *accuracy* or *correctness*.

An independently constructed phonemic inventory therefore provides a snapshot of the child's phonological knowledge; however, a child's inventory can also be compared to the inventory of the target adult language to provide additional information. This comparison generates a relational measure of those phonemes that are *missing* from the individual's phonemic inventory. Thus, an inventory of the segments that are 'in' a child's phonemic inventory is an independent measure, and an inventory of the phonemes that are 'missing' from a child's phonemic inventory is a relational (i.e. comparative) measure because labelling phonemes as 'missing' requires a comparison to the adult target inventory (Dinnsen, 1984). A description of the phonemes that are either 'in' or 'missing' from the inventory provides simultaneously a neutral snapshot of a child's phonological system and a comparative indication of the weaknesses or gaps in phonological knowledge. This information is used to determine the presence or severity of PD (e.g. Gierut et al., 1994), monitor change over time (e.g. Gierut, 1992), and to guide the selection of appropriate speech-sound targets and goals for intervention (e.g. Barlow & Gierut, 2002; Morrisette, Farris, & Gierut, 2006).

Despite its unique informativeness, phonemic inventory analysis is not commonly employed by practicing SLPs (McLeod & Baker, 2014). Perhaps this is due to the abstract nature of the knowledge it captures or the opacity of the underlying generative assumptions from which this measure is derived, but there are clear logistical barriers as well. Ferguson and Farwell (1975) and Gierut et al. (1994) describe these obstacles to phonemic inventory analysis in children, including the variability of their word productions and the difficulty of obtaining sufficient words to serve as minimal pairs. Certainly, the descriptive process required to identify minimal pairs and generate a child's phonemic inventory requires collection of a thorough speech sample strategically designed to capture contrastive minimal pairs, which may be time-prohibitive for many practicing clinicians.

Production accuracy

Whereas phonemic inventories provide insight into a child's functional phonological knowledge, other frequently employed speech sound measures eschew underlying knowledge and instead capture production accuracy. A consonant accuracy measure compares each consonant segment produced by the child to its corresponding target (i.e. adult-like) form. This pairwise comparison requires no assumption of underlying phonological function, and it generates a percentage accuracy score that is relational, quantitative, and immediately interpretable. The most commonly used segmental accuracy measure is Percentage of Consonants Correct-Revised (PCC-R; Shriberg, Austin, Lewis, McSweeny, & Wilson, 1997), which collapses across all consonants to provide a single accuracy percentage from a given sample. However, an SLP may also choose to examine consonant accuracy for each consonant separately to provide more nuanced accuracy information. Furthermore, with the advent of computer-assisted analysis, including freely available software (e.g. Phon; Rose & Hedlund, 2017), SLPs can calculate consonant accuracy measures consistently and relatively quickly (Byun & Rose, 2016).

Despite their differences, phonological knowledge and accurate production are presumed to be related to one another and are often discussed jointly (e.g. Gierut, Elbert, & Dinnsen, 1987). However, it is notable that measures of one can contrast with the other. One such instance is discussed in Dinnsen and Barlow (1998) and Dinnsen, Green, Gierut, and Morrisette (2011). In their example, a child uses the consonant /θ/ phonemically to contrast words, yet this same child never uses /θ/ accurately due to a chain-shift substitution pattern, such that [θ] is produced exclusively as a substitution for /s/ (i.e. dentalisation), and every instance of target /θ/ is produced as [f] (i.e. labialisation). These patterns result in productions, such as [fʌm] for 'thumb' and [θʌm] for 'some', which serve as a minimal pair for both /f/ and /θ/. Consequently, /θ/ would be considered phonemic and thus 'in' the child's phonemic inventory, despite the child's 0% accuracy for production of /θ/. Given the potential for divergence, there is motivation to better evaluate the relationship between phonemic inventory (a measure of phonological knowledge) and consonant accuracy (a measure of adult-like production).

Current study

To better understand these different phonological assessment measures, the purpose of the investigation described here is twofold. Our first goal is to identify the relationship between phonemic inventory – a qualitative, linguistically motivated measure of phonological knowledge – with consonant accuracy – a quantitative, relational measure that directly compares a child's phonological productions to the target form. By identifying a relationship between these two measures, we improve our understanding of how production accuracy reflects phonological knowledge in children with PD. Our second goal is to determine if the relationship between the measures would permit identification of a percentage accuracy cutoff score (or cutoff range) that reliably classifies sounds as 'in' or 'out' of a child's phonemic inventory in speech-sound probes of varying length. A percentage accuracy cutoff suggestive of the phonemic function of a given consonant could provide useful information about phonological knowledge without the time-consuming process of identifying minimal pair contrasts.

Method

Participants

Data for this study were drawn from 275 children between 3 and 8.5 years old (mean age = 4;4), whose single-word productions were transcribed as part of the Learnability Project (Gierut, 2015b). Participants in the Learnability Project were monolingual, English-speaking children residing in the Midwestern United States who presented with functional PD, determined by performance > 1 SD below the mean on the first or second edition of the Goldman-Fristoe Test of Articulation (GFTA/GFTA-2; Goldman & Fristoe, 1986, 2000) and a reduced phonemic inventory, missing at least 6 target English consonants. Furthermore, all participants had normal hearing, no documented history of motor or otherwise organic disorders, no indication of cognitive delay, and normal oral-motor function. Participating children received experimental speech intervention; however, data in this study come only from the children's pre-treatment samples. Additional demographic information for participants in the Learnability Project can be found in Gierut (2015b).[1]

[1]Archival data were retrieved from the Gierut / Learnability Project collection of the IUScholarWorks repository at https://scholarworks.iu.edu/dspace/handle/2022/20061 The archival data were original to the Learnability Project and supported by grants from the National Institutes of Health to Indiana University (DC00433, RR7031K, DC00076, DC001694; PI: Gierut). The views expressed herein do not represent those of the National Institutes of Health, Indiana University, or the Learnability Project. The author(s) assume(s) sole responsibility for any errors, modifications, misapplications, or misinterpretations that may have been introduced in extraction or use of the archival data.

Data transformation

Data were phonetic transcriptions of children's productions of words in the Phonological Knowledge Probe (PKP; Gierut, 1986), a single-word probe eliciting 293 words, and the GFTA or GFTA-2, eliciting 44 and 53 words, respectively. Original archival transcriptions included the orthography of the target word and transcription of the child's production in IPA notation. Reliability for 10% of archival consonant transcriptions was reported at 93% (Gierut, 2015a).

To facilitate analyses using Phon (v2.2; Rose & Hedlund, 2017), data were translated from their archival Excel format to Phon-readable Unicode text using a Python script. Non-standard notation conventions were translated to standard IPA notations compatible with Phon. For instance, the US English rhotic consonant, transcribed as [r] in the archival data, was translated to the standard IPA notation [ɹ]. Some diacritic symbols in the original data, such as [^β], were not available as characters in Phon. In these instances, the symbol was changed to a similar diacritic (e.g. [ᵇ]) and appended with [ˤ] (e.g. [deɪ^β] became [deɪᵇˤ]) to document the change during translation. These diacritic differences between the original archival transcription and the translated format did not impact our analyses, as these changes were implemented consistently across the data. Furthermore, diacritic symbols were ignored during consonant accuracy calculations, as described below.

Orthographic and IPA transcriptions translated directly from archival data were sufficient for extracting each child's phonemic inventory, as this measure did not require comparison to the target production of each word. However, calculation of consonant accuracy required transcription of the target, adult-like form of each word. Given the scope of the data to be analysed (approximately 93,000 words or 243,000 consonants), we generated a single, representative set of target transcriptions for all sampled word productions to permit relational analyses. A two-step process generated these target transcriptions for comparison to the children's productions. First, broad target transcriptions were generated for each word in the PKP, GFTA, and GFTA-2 from the English IPA dictionary in Phon. Second, two research assistants (undergraduate and graduate students of speech-language pathology or linguistics) reviewed archival transcriptions extracted from 200 participants and compared these to the dictionary-generated targets to arrive at consensus for a single target transcription for each word deemed to best capture the dialect spoken by these children and the transcription conventions of the archival data. These transcriptions were also reviewed by the first author. Once confirmed, these target transcriptions were aligned to each child's transcribed productions using the English syllabification and alignment algorithms in Phon and then compared to generate the relational consonant accuracy measure used in this study.

To validate the generated target forms, alignment, and our automated percentage consonant accuracy measure, percentage consonant accuracy for word productions in the PKP was calculated manually by research assistants for 20% of participants. Procedures for manual calculation followed those outlined for calculation of PCC-R. On average, manually calculated accuracy deviated 3.9% (SD = 3.7%) from automatically generated values. Correlation between manually derived accuracy and automated accuracy measures was 0.96.

Variables

Two primary measures were derived for each of 23 American English consonants (excluding /ʒ/ due to limited sampling of this consonant), for each child. The first measure was a binary, categorical designation of phoneme status. For a given child, if two contrastive minimal pairs were identified for a given consonant, that consonant was deemed phonemic and coded as 'in' the phonemic inventory. For many children, one or more non-ambient sounds (e.g. /wʴ/ or /ʔ/) were also used contrastively; however, only the phonemic status of ambient phonemes was recorded because the corresponding accuracy measure is only derivable for ambient consonants. When two minimal pairs were not identified, that consonant was coded as 'out' of the child's phonemic inventory. Minimal pairs were identified, and phoneme status was confirmed using the AutoPATT plugin for Phon (Combiths, Amberg, & Barlow, 2016). Data used to calculate this measure were participants' word productions from the 293-item PKP (rather than the shorter GFTA or GFTA-2) to obtain sufficient opportunities for two contrastive minimal pairs for each of 23 English phonemes.

The second measure was a quantitative measure of consonant production accuracy. For a given child, accuracy was calculated for their production of each English consonant in Phon by comparing each target consonant with its corresponding segment in the child's production. This comparison was automated with a consonant accuracy query in Phon. In order to provide an accuracy calculation that is easily replicable and robust to varied ages and severities of impairment, we followed the same procedures used for the global measure of PCC-R (Shriberg et al., 1997). Unlike the original Percentage of Consonants Correct measure (Shriberg & Kwiatkowski, 1982), PCC-R ignores distortions in its calculation. Although distortion patterns are diagnostically informative, their absence from these calculations make the measure simpler, less prone to error, and appropriate for a more diverse population of children. Furthermore, PCC-R is a well-attested and reliable consonant accuracy measure (see Shriberg et al., 1997). Following PCC-R procedures, to be coded as correct, the child's production was required only to match the base target phone. For instance, production of [s̪] for target /s/ was considered correct; however, phonemic substitution, such as [f] for target /θ/, or omission of a target consonant were considered incorrect. By these criteria, each child was designated a percentage accuracy for each of the 23 English consonants. Because consonant accuracy may be more robust to varying sample length than phonemic inventory (which requires multiple minimal pair opportunities for each phoneme), the accuracy measure was calculated separately for productions in the PKP and the GFTA/GFTA-2. This permitted comparison between probe types.

Additional variables used in the analyses were participant age, sample type (PKP, GFTA/GFTA-2), and normative age of acquisition (early, middle, late) for each English consonant, as categorized in (Shriberg, 1993).

Analyses

Logistic regression determined the ability of percentage consonant accuracy to predict the phonemic inventory measure, including the mitigating impacts of sample length, child age, and consonant age of acquisition. Receiver-operating characteristic (ROC) curve analysis determined cutoff accuracy values with optimal sensitivity and specificity according to elicited sample length, child age, and age of acquisition. Regression models, ROC curves, and optimal cutoff value estimation were conducted in R (R Core Team, 2013) using pROC, OptimalCutpoints, and visreg packages.

Results

Descriptive statistics

Mean percentage accuracy for all PKP consonants 'out' of the children's phonemic inventories was 12.0% SD 23.3%. Mean percentage accuracy for all PKP consonants 'in' the children's phonemic inventories was 74.0% SD 26.4%. Mean accuracy for 'out' GFTA/GFTA-2 consonants was 12.1% SD 25.0%. Mean accuracy for 'in' GFTA/GFTA-2 consonants was 70.9% SD 31.1%. Thus, phonemic consonants were produced with greater accuracy than non-phonemic consonants, and mean accuracies for phonemic and non-phonemic consonants were similar across probes. Means and standard deviations for PKP consonants according to phonemic inventory classification, normative age of acquisition, and participant age are displayed in Table 1.

Logistic regression

Phoneme status of 23 American English consonants, excluding /ʒ/, was predicted by consonant accuracy ($p < 0.01$) and classification as an early-, middle-, or late-acquired consonant ($p < 0.01$). As expected, consonants with higher accuracy and those that are earlier-acquired are more likely to be used as phonemes by the child. The main effect of child age on phoneme status was not significant ($p = 0.33$). Significant Consonant Accuracy × Child Age ($p < 0.01$) and Consonant Accuracy × Age of Acquisition ($p < 0.01$) interactions also emerged, such that consonant accuracy was most predictive of phoneme status in younger children and for middle- and late-acquired consonants. These interactions are displayed in Figure 1.

Receiver operating characteristic curve analysis

The ability of a percentage consonant accuracy cutoff to classify the phoneme status of 23 English consonants was quantified with ROC curve analysis, using several paradigms to determine the optimum cutoff value. For the larger 293-item PKP, a consonant

Table 1. Mean consonant accuracy by phonemic inventory, age of acquisition, and participant age.

	'Out' consonants	*'In' phonemes*
	Age of acquisition	
Early	52.8% (35.6%)	83.9% (19.4%)
Middle	8.4% (17.6%)	68.0% (26.3%)
Late	9.2% (19.3%)	55.6% (30.1%)
	Participant age	
<5 Years	11.0% (21.9%)	72.8% (26.7%)
≥5 Years	16.6% (28.3%)	78.3% (24.7%)

Note. Consonant age of acquisition classification based on Shriberg (1993). Standard deviations in parentheses.

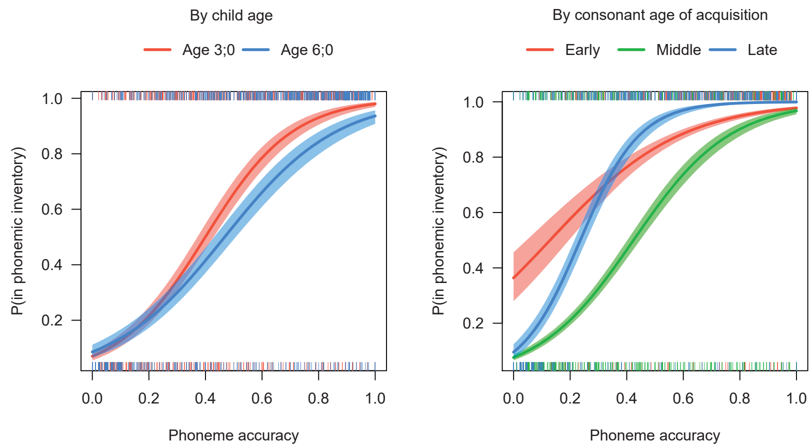


Figure 1. Phonemic inventory classification.

Note. Regression trends are graphed on a logarithmic scale, displaying the probability of phonemic inventory inclusion according to phoneme accuracy. To illustrate the Consonant Accuracy × Child Age interaction, data are shown for children at 3 and 6 years of age (reflective of preschool and school-age groups, respectively). Steeper curves indicate better predictive power. Ticks along the upper and lower plot borders indicate actual data points.

accuracy of 20.4% was the most efficient cutoff, correctly classifying the phoneme status of 90.0% of English consonants for all 275 children (sensitivity = 94.9%; specificity = 83.1%). Other potential percentage accuracy cutoff values, derived from various methods for determination of optimal classification, including maximum efficiency (i.e. most accurate classification; Galen, 1986; Greiner, 1996), Youden's Index (Greiner, Pfeiffer, & Smith, 2000; Youden, 1950), and closest to ROC plot point 0,1 (Metz, 1978; Vermont et al., 1991), are displayed in Table 2.

Furthermore, sample length impacted potential cutoff score classification accuracy. Classification accuracy, sensitivity, and specificity of several potential consonant accuracy cutoff values are displayed for data from the 293-item PKP and the 44–53-item GFTA/GFTA-2 in Table 3. When data are drawn from a larger (ostensibly more thorough) sample, optimal consonant accuracy cutoff values are lower, and sensitivity, specificity, and classification accuracy are higher than when data are drawn from a smaller sample. Accordingly, the area under the receiver operating characteristic curve (AUC/AUROC) was higher for the PKP data (AUROC = 0.94, 95% CI = [0.934, 0.947]) than for the GFTA/GFTA-2 data (AUROC = 0.90, 95% CI = 0.895, 0.911). Note that an AUROC closer to 1 has better overall classification ability. ROC curves for phoneme classification based on

Table 2. Potential consonant accuracy cutoff values predictive of phoneme status.

	Maximum Efficiency	Youden's Index	Closest to ROC (0,1)
Cutoff	20.4%	21.1%	30.2%
Class.	90.1%	90.0%	89.1%
Sens.	94.9%	83.6%	90.9%
Spec.	83.1%	88.9%	86.6%

Note. Cutoff = optimum percent accuracy cutoff value. Class. = classification accuracy. Sens. = sensitivity. Spec. = specificity.

Table 3. Phoneme classification metrics for potential accuracy cutoff values by sample type.

	20% Cutoff			30% Cutoff			40% Cutoff			50% Cutoff		
	Sens.	*Spec.*	*Class.*	*Sens.*	*Spec.*	*Class.*	*Sens.*	*Spec.*	*Class.*	*Sens.*	*Spec.*	*Class.*
PKP	0.95	0.83	0.90	0.91	0.87	0.89	0.85	0.89	0.87	0.79	0.91	0.84
GFTA	0.90	0.80	0.86	0.86	0.83	0.85	0.81	0.88	0.84	0.73	0.92	0.80

Note. Sens. = sensitivity. Spec. = specificity. Class. = classification accuracy.

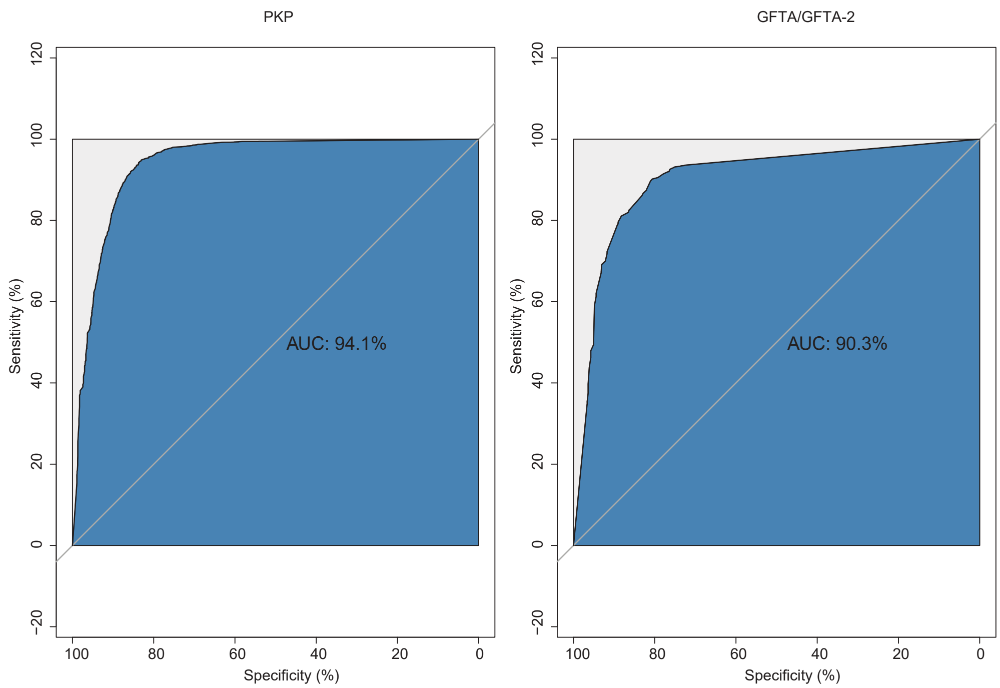


Figure 2. ROC curve for phoneme classification via consonant accuracy by sample type.

consonant accuracy data from the PKP and GFTA/GFTA-2 are displayed in Figure 2. These findings are discussed in the next section.

Discussion

In this study, a qualitative measure of phonological knowledge and function (i.e. phonemic inventory) was compared to a quantitative accuracy measure (i.e. percentage consonant accuracy) in young, monolingual English-speaking children with PD. A strong relationship emerged between a given consonant's percentage accuracy and its contrastive, phonemic use. Furthermore, ROC curve analyses indicated that a relatively low consonant accuracy cutoff (approximately 20–30%) can correctly classify up to 90% of English consonants as either 'in' or 'out' of a given child's phonemic inventory using data from the 293-item PKP. In other words, when a child produces a given English consonant with a percentage accuracy above the cutoff of 20–30%, it is more likely that this child already uses the consonant phonemically to contrast words. Our analyses also found that sensitivity was higher than specificity for this cutoff range, indicating that a percentage accuracy cutoff is better at correctly including phonemic consonants in the inventory than correctly excluding non-phonemic consonants. Finally, the ability of percentage consonant accuracy to predict a child's phonemic use of a given consonant in these data was also mitigated by several factors, including speech sample length, child age, and normative age of acquisition of the consonant in question.

A percentage consonant accuracy cutoff was poorer at classifying a child's phonemic inventory when consonant accuracy was derived from the GFTA/GFTA-2 (i.e. a short sample of 44–53 words). The highest classification accuracy of 90% was only achievable with productions from the PKP, a larger speech sample of 293 words. In addition to sampling size, differences in the predictive power of consonant accuracy across the two probes may also have been related to qualitative differences between them. The PKP was designed to capture phonological knowledge in a research context and, consequently, was constructed with many production opportunities for each consonant (at least 5 in each word position) and to allow opportunities to demonstrate contrastive minimal pairs (Gierut, 2015c). Conversely, the GFTA/GFTA-2 is a standardized testing instrument designed for rapid administration. In typical usage, an examiner derives a single score, collapsed across all consonants, for comparison to a normative database. Although the GFTA/GFTA-2 is markedly shorter than the PKP, it is also likely that differences in the depth and breadth of these samples contributed to differences in their ability to provide accuracy calculations sufficient to predict a consonant's phonemic status.

Finally, percentage consonant accuracy was also better able to predict the phonemic status of consonants produced by younger, preschool-age children than those produced by older, school-age children, and phonemic classification based on consonant accuracy was more accurate for normatively late-acquired sounds (/ʃ, s, θ, ð, ɹ, z, l/) than for normatively middle-acquired (/t, ŋ, k, ɡ, f, v, ʧ, ʤ/) and especially for early-acquired sounds (/m, b, j, n, w, d, p, h/; Shriberg, 1993). Older children are more advanced developmentally and, thus, less likely to demonstrate variable accuracy rates and phoneme usage. Similarly, earlier-acquired consonants are more likely to be produced accurately and used phonemically. These ceiling effects could be applicable to the broader population of children with PD, but they are also likely confounded by distributional limitations of the study sample, and this will be discussed more below.

Implications for assessment

The relationship between consonant accuracy and phonemic inventory identified in this study has potential implications for phonological assessment. By describing and quantifying the relationship between a qualitative, descriptive measure of phonological knowledge and a quantitative, relational measure of accurate production, we confirm the informativeness of both measures and highlight similarities between them. Although seemingly a simple comparison between two measures of speech-sound usage, phonemic inventory and percentage consonant accuracy represent divergent conceptualizations of phonological skill. A phonemic inventory is intended to describe a child's contrastive speech-sound units (i.e. phonemes), and, as such, several generativist linguistic assumptions are required for meaningful interpretation of this measure. The term 'phonemic inventory' was first derived from linguistic descriptions of fully formed adult languages spoken by an entire community of speakers. The extension of this measure to the analysis of child phonology borrows from the descriptivist tradition, requiring instances of minimal pairs to confirm the contrastive role of a given speech-sound phoneme. On the other hand, consonant accuracy does not require any assumption of a consonant's underlying phonemic function. Rather, it relies on a direct comparison of each consonant in the child's word productions to its corresponding target consonant in the adult-like form of the intended word. The higher a child's percentage accuracy for a given consonant, the more closely the child's production of that consonant coincides with target, adult productions.

Despite the inherent differences between these two measures, the findings of this study confirm that there is considerable overlap in the useful information they capture. Further still, the quantitative, relational information provided by percentage consonant accuracy may provide a relatively accurate estimate of a preschool-age child's functional phonemic usage of later-developing consonants, but only when this accuracy measure is derived from a sufficiently thorough sample. Consequently, the relationship between measures of qualitative phonological knowledge and quantitative production accuracy should be considered in assessment and subsequent treatment goal selection and progress monitoring – especially given the increasing availability of (often quantitative) computer-assisted measures that have the potential to shift the assessment landscape.

Limitations and future directions

Limitations in the archival data used in this study likely contributed to the poorer predictive power of consonant accuracy for earlier acquired sounds and older children. The participants in this study, as expected, demonstrated greater mastery of early-acquired consonants, as indicated by less variable, generally higher accuracy rates and more frequent inclusion in their phonemic inventories. This ceiling effect likely impacted the ability to predict phonemic inventory inclusion from consonant accuracy for these consonants. It is possible that phonemic use of early-acquired sounds could be predictable from consonant accuracy given data with more variable early-acquired consonant accuracy rates and phonemic use, such as with younger children or those with more severe impairment. The poorer predictive power of consonant accuracy to categorize phonemic function for school-age children is also likely impacted by the participants' age distribution in these data. The majority of 275 participants were preschool-aged, with only 57 school-aged children in the sample. Consequently, poorer predictions for school-aged children may simply reflect the limited sampling of school-aged children in these data.

Future work examining the relationship between measures of phonemic function and quantitative accuracy could address these sampling limitations through prospective data collection involving younger children or those with more severely impacted phonological systems as well as a greater number of school-aged children. Although the role of child age and normative acquisition trajectories require further investigation, the relationship between phonemic inventory inclusion and consonant accuracy identified in the current data remain most robust for late-acquired sounds in preschool-aged children with PD.

Finally, the lower sensitivity of percentage accuracy in determining phonemic status suggests that, when classification error does occur, it is more likely to result in over-identification of phonemic consonants. Future work should identify and compare the clinical impact of over- and under-estimation of phonological knowledge to determine which of these error types is most important to minimize. This type of work could guide modification of optimal percentage accuracy cutoff values or the development of other criteria to improve the clinical utility of estimates of phonological knowledge. As our understanding of the relationship between phonemic inventory and consonant accuracy measures improves, clinicians may eventually be able to infer information about functional phonological knowledge from a quantitative accuracy measure, which could streamline assessment, treatment target selection, and progress monitoring for children with PD.

Disclosure Statement

A portion of this research was funded by NIH NIDCD DC017697-01. The authors have no other declarations of interest to report.

Funding

This work was supported by the National Institute on Deafness and Other Communication Disorders [DC017697-01].

ORCID

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

References

Barlow, J. A., & Gierut, J. A. (2002). Minimal pair approaches to phonological remediation. *Seminars in Speech and Language*, *23*, 57–68. doi:10.1055/s-2002-24969

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., Inglis, A., & Lancee, W. (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(6), 815–825. doi:10.1097/00004583-199606000-00022

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(6), 804–814. doi:10.1097/00004583-199606000-00021

Byun, T. M. A., & Rose, Y. (2016). Analyzing clinical phonological data using Phon. *Seminars in Speech and Language*, *37*(2), 85–105. doi:10.1055/s-0036-1580741

Chomsky, N., & Halle, M. (1968). *The sound pattern of english*. New York, NY: Harper & Row.

Combiths, P. N., Amberg, R., & Barlow, J. A. (2016). *Phonological analysis software in assessment and treatment of speech-sound disorders*. Paper presented at the Convention of the American Speech-Language-Hearing Association, Philadelphia, PA.

Dinnsen, D. A. (1984). Methods and empirical issues in analyzing functional misarticulation. In M. Elbert, D. A. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulating child (ASHA Monographs No. 22)* (pp. 5–17). Rockville, MD: ASHA.

Dinnsen, D. A., & Barlow, J. A. (1998). On the characterization of a chain shift in normal and delayed phonological acquisition. *Journal of Child Language*, *25*, 61–94.

Dinnsen, D. A., Green, C. R., Gierut, J. A., & Morrisette, M. L. (2011). On the anatomy of a chain shift. *Journal of Linguistics*, *47*(2), 275–299. doi:10.1017/S0022226710000368
Felsenfeld, S., Broen, P. A., & McGue, M. (1992). A 28-year follow-up of adults with a history of moderate phonological disorder: Linguistic and personality results. *Journal of Speech and Hearing Research*, *35*, 1114–1125.
Felsenfeld, S., Broen, P. A., & McGue, M. (1994). A 28-year follow-up of adults with a history of moderate phonological disorder: Educational and occupational results. *Journal of Speech and Hearing Research*, *37*, 1341–1353.
Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition: English initial consonants in the first fifty words. *Language*, *51*, 419–439. doi:10.2307/412864
Fry, D. B. (1967). The phonemic system in children's speech. *British Journal of Disorders of Communication*, *3*(1), 13–19. doi:10.3109/13682826809011436
Galen, R. S. (1986). Use of predictive value theory in clinical immunology. In *Manual of clinical laboratory immunology* (3rd ed., pp. 966–970). Washington, DC: American Society for Microbiology.
Gierut, J. A. (1986). *On the relationship between phonological knowledge and generalization learning in misarticulating children* (Doctoral Dissertation, 1985). Indiana University Linguistics Club, Bloomington.
Gierut, J. A. (1992). The conditions and course of clinically induced phonological change. *Journal of Speech and Hearing Research*, *35*, 1049–1063.
Gierut, J. A. (2015a). Experimental designs and protocols. *Learnability Project Working Paper*. Bloomington: Indiana University.
Gierut, J. A. (2015b). Participant eligibility and demographics. *Learnability Project Working Paper*. Bloomington: Indiana University.
Gierut, J. A. (2015c). Phonological protocols. *Learnability Project Working Paper*. Bloomington: Indiana University.
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*, 462–479.
Gierut, J. A., Simmerman, C. L., & Neumann, H. J. (1994). Phonemic structures of delayed phonological systems. *Journal of Child Language*, *21*, 291–316.
Goldman, R., & Fristoe, M. (1986). *Goldman-fristoe test of articulation*. Circles Pines, MN: American Guidance Service.
Goldman, R., & Fristoe, M. (2000). *Goldman-fristoe test of articulation-2*. Circle Pines, MN: American Guidance Service, Inc.
Greiner, M. (1996). Two-graph receiver operating characteristic (TG-ROC): Update version supports optimisation of cut-off values that minimise overall misclassification costs. *Journal of Immunological Methods*, *191*, 93–94.
Greiner, M., Pfeiffer, D., & Smith, R. D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventive Veterinary Medicine*, *45* (1–2), 23–41.
Jakobson, R. (1968). *Child language, aphasia and phonological universals*. The Hague, Netherlands: Mouton.
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. doi:10.1080/136828200247133
Lewis, B. A., Patton, E., Freebairn, L., Tag, J., Iyengar, S. K., Stein, C. M., & Taylor, H. G. (2016). Psychosocial co-morbidities in adolescents and adults with histories of communication disorders. *Journal of Communication Disorders*, *61*, 60–70. doi:10.1016/j.jcomdis.2016.03.004
McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics*, *28*(7–8), 508–531. doi:10.3109/02699206.2014.926994

Metz, C. E. (1978). Basic principles of ROC analysis. *Seminars in Nuclear Medicine*, *8*(4), 283–298.
Morrisette, M. L., Farris, A. W., & Gierut, J. A. (2006). Applications of learnability theory to clinical phonology. *Advances in Speech-Language Pathology*, *8*, 207–219. doi:10.1080/14417040600823284
Peterson, R. L., Pennington, B. F., Shriberg, L. D., & Boada, R. (2009). What influences literacy outcome in children with speech sound disorder? *Journal of Speech, Language, and Hearing Research*, *52*(5), 1175–1188. doi:10.1044/1092-4388(2009/08-0024)
R Core Team. (2013). *R: A language and environment for statistical computing [Computer software]*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/
Rose, Y., & Hedlund, G. (2017). *Phon (Version 2.2) [Computer software]*. Retrieved from https://www.phon.ca/phontrac
Shriberg, L. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research*, *36*, 105–140.
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*, 708–722. doi:10.1044/jslhr.4004.708
Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*, *47*, 256–270.
Vermont, J., Bosson, J., Francois, P., Robert, C., Rueff, A., & Demongeot, J. (1991). Strategies for graphical threshold determination. *Computer Methods and Programs in Biomedicine*, *35*(2), 141–150.
Voegelin, C. F. (1957). Six statements for a phonemic inventory. *Journal of American Linguistics*, *23*(2), 78–84. doi:10.1086/464396
Youden, W. J. (1950). Index for rating diagnostic tests. *Cancer*, *3*(1), 32–35.

Chapter 3, in full, is a reprint of material as it appears in Combiths, P., Barlow, J., & Sanchez, E., (2019). Quantifying phonological knowledge in children with phonological disorder. *Clinical Linguistics & Phonetics, 33*(10–11), 885–898. The dissertation author was the primary investigator and author of this paper.

CHAPTER 4:

Influences of Phonological Context on Tense Marking in Spanish–English Dual Language Learners

JSLHR

Research Article

Influences of Phonological Context on Tense Marking in Spanish–English Dual Language Learners

Philip N. Combiths,[a,b] Jessica A. Barlow,[a] Irina Potapova,[a,b] and Sonja Pruitt-Lord[a]

Purpose: The emergence of tense-morpheme marking during language acquisition is highly variable, which confounds the use of tense marking as a diagnostic indicator of language impairment in linguistically diverse populations. In this study, we seek to better understand tense-marking patterns in young bilingual children by comparing phonological influences on marking of 2 word-final tense morphemes.

Method: In spontaneous connected speech samples from 10 Spanish–English dual language learners aged 56–66 months (M = 61.7, SD = 3.4), we examined marking rates of past tense *-ed* and third person singular *-s* morphemes in different environments, using multiple measures of phonological context.

Results: Both morphemes were found to exhibit notably contrastive marking patterns in some contexts. Each was most sensitive to a different combination of phonological influences in the verb stem and the following word.

Conclusions: These findings extend existing evidence from monolingual speakers for the influence of word-final phonological context on morpheme production to a bilingual population. Further, novel findings not yet attested in previous research support an expanded consideration of phonological context in clinical decision making and future research related to word-final morphology.

Children's acquisition of grammatical morphemes can be characterized as an initial absence of the morpheme that is followed by a period of inconsistent use, during which the morpheme is considered to be emerging. Mastery of the morpheme occurs once the morpheme is used correctly and consistently, with common benchmarks for achievement of mastery falling between 80% and 95% accuracy (e.g., Brown, 1973). Children with typical development from all linguistic backgrounds can be expected to move reliably from absence to mastery; however, the mechanisms underlying this trajectory are not well understood.

This study is concerned with the emerging stage of morpheme acquisition. Because this stage is highly variable and can appear differently across populations, including in children with language impairment (LI; e.g., Gutiérrez-Clellen, Simon-Cereijido, & Wagner, 2008; Paradis, 2005; Paradis & Crago, 2000; Rice, Tomblin, Hoffman, Richman, & Marquis, 2004), there is considerable motivation to understand the factors that affect morpheme production during this period. Understanding patterns of influence in different population groups is of scientific interest because it contributes to the growing body of typological evidence for factors that interact with morpheme production. Further, influences on morpheme production are of clinical interest for their diagnostic potential, particularly in differentiating nonclinical language differences, such as characteristics of a nonmainstream dialect or multilingual speaker, from a language disorder (Paradis, 2005).

Word-Final Tense Morphology in Linguistically Diverse Populations

In English, regular tense marking is mostly restricted to two suffixes that attach to the right edge of the verb stem (i.e., word-final morphemes): past tense (PT; e.g., *-ed* in *walked*) and third person singular present tense (3s; e.g., *-s* in *plays*). In particular, children with LI demonstrate distinct error patterns and protracted emergence of PT and 3s morphemes, resulting in an extended period of lower marking rates when compared with age- or language-matched peers (for an overview, see Leonard, 2014). Children's

[a]San Diego State University, CA
[b]University of California, San Diego

Correspondence to Philip N. Combiths: pcombiths@sdsu.edu

Editor: Sean Redmond
Associate Editor: Emma Hayiou-Thomas

Received October 19, 2016
Revision received February 3, 2017
Accepted March 10, 2017
https://doi.org/10.1044/2017_JSLHR-L-16-0402

Disclosure: The authors have declared that no competing interests existed at the time of publication.

productive use of tense morphology, such as PT and 3s, has therefore emerged as a particularly salient indicator of LI in monolingual English speakers (e.g., Bedore & Leonard, 1998; Rice & Wexler, 1996).

Although the diagnostic value of tense marking has been established for monolingual English speakers, the validity of its use among bilingual speakers, particularly young dual language learners (DLLs), has been called into question. These DLLs are developing two languages at potentially incongruous stages of linguistic development. This is particularly true for sequentially bilingual children exposed first to their native language and later to a second language (for further discussion of this terminology, see Barlow & Enríquez, 2007; Hammer, Miccio, & Rodriguez, 2004). When the later-acquired second language is English, performance on measures of English is highly variable across individuals and may be lower than that of their same-age monolingual peers (Iglesias & Rojas, 2012).

This performance variability also extends to tense-morpheme marking, which is especially relevant in that several studies have shown tense-morpheme marking rates among bilingual DLLs without LI to be similar to or even lower than rates for age-matched monolingual speakers with LI (e.g., Gutiérrez-Clellen et al., 2008; Paradis & Crago, 2000; Paradis, Rice, Crago, & Marquis, 2008). However, this potentially confounding overlap has been identified as a result of comparing bilingual groups to monolingual speakers with LI (i.e., across linguistic populations). When compared within a bilingual group matched by dual-language profile, tense-morpheme use has been shown to differentiate children with and without LI (Blom & Paradis, 2013; Gutiérrez-Clellen et al., 2008; Jacobson & Schwartz, 2005).

In short, comparing morpheme marking rates across monolingual and bilingual populations can result in under- or overidentification of LI (Paradis, 2005). However, the feasibility of using tense-morpheme marking to identify LI within bilingual and monolingual groups suggests that tense morphology is related to LI in both, and identification of LI using morpheme-marking differences is still plausible, even within linguistically diverse populations. Though comparison of tense-morpheme use across linguistic groups is currently problematic, this may be due to the broad lens with which we have so far examined morpheme marking rates. It could be that examining other linguistic influences on morpheme marking would allow us to identify patterns in LI groups that are sufficiently robust to identify LI not only within but also across linguistically diverse populations.

Influences on Tense-Morpheme Production

In order to make use of the differentiating power of tense-morpheme marking in other populations, such as DLLs, a more complete understanding of the factors that affect marking of these morphemes is needed. Morpheme production, as with other areas of language, does not exist in isolation from other influences. To date, multiple factors related to the context in which the morpheme occurs have been shown to influence emerging PT and 3s marking rates. These include nonlinguistic influences, such as task type (Barlow, Pruitt-Lord, & Combiths, 2015; Oetting et al., 2012); supralexical linguistic influences, such as utterance position (Barlow & Pruitt-Lord, 2014; Dalal & Loeb, 2005; Song, Sundara, & Demuth, 2009; Sundara, Demuth, & Kuhl, 2011); and lexical influences, such as the frequency with which verbs are inflected for tense marking (Blom & Paradis, 2013; Blom, Paradis, & Duncan, 2012; Marchman, 1997; Marchman, Wulfeck, & Weismer, 1999; Oetting & Horohov, 1997).

In this study, we explored sublexical influences on morpheme production, particularly the influence of phonological context. There is precedent for the influence of phonology—particularly of the preceding segment in the verb stem—on morpheme marking in acquisition and dialect patterns. Phonological context has been shown to affect regular PT and 3s marking rates across studies of groups both with and without impairment, spanning linguistic populations and task types (e.g., Barlow & Pruitt-Lord, 2014; Johnson & Morris, 2007). However, indices used to measure phonological-context effects have been inconsistent, and findings even appear contradictory in some cases. See Table 1 for a summary of previous findings, which will be discussed later.

PT Marking and Word-Final Phonological Context

In terms of PT marking, consideration of phonological context has mostly been localized to the end of the root verb, particularly the root-final segment, which immediately precedes the morpheme segment (e.g., the root-final segment /p/ in *drop* immediately precedes [-t] in *dropped*). Initial studies of the influence of phonological context on morpheme marking rates directly examined the influence of preceding phonological contexts that result in either monosegmental or syllabic allomorphs (i.e., [-t] or [-d], as in *popped* /pɑpt/ or *filled* /fɪld/, vs. syllabic [-ɪd], as in *started* /stɑɹtɪd/), converging in their findings that PT is marked more frequently in contexts that elicit the monosegmental allomorph than those that elicit the syllabic form (Berko, 1958; Blom & Paradis, 2013; Marchman, 1997; Marchman et al., 1999).

Sequential complexity. The heterogeneity with which phonological context has been operationalized in previous research complicates a cohesive discussion of previous findings. In the interest of a unified discussion of phonological context, we appeal to the construct of phonological complexity, which provides a useful framework for contextualizing various aspects of phonology. The relative simplicity or complexity of a phonological unit or sequence is based on many factors, including universal preferences for certain structures across the world's languages (Jakobson, 1968), order of acquisition in languages that contain these structures, and simplification patterns in developing or disordered speech (Barlow & Gierut, 1999; Hawkins, 1987). For instance, the sequential arrangement of speech sound types contributes to the relative complexity of a word or phrase. To be specific, adjacent consonants, also known as consonant clusters (CC or CCC), are more complex (i.e.,

Table 1. Abbreviated findings for word-final phonological-context effects on regular marking of past tense and third person singular.

Study	LI/TD	Linguistic population	Phonological context	Task	Observed effect[a]
Past tense					
Berko (1958)	TD	MAE	Allomorphic	Elicitation	Segmental > syllabic
Oetting & Horohov (1997)	Both	MAE	Son vs. obs	Elicitation	Son > obs
Marchman (1997)	TD	MAE	Allomorphic/probability	Elicitation	Segmental > syllabic/high > low prob
Marchman et al. (1999)	Both	MAE	Allomorphic	Elicitation	Segmental > syllabic
Johnson & Morris (2007)	Both	MAE	Son vs. obs	Imitation	Son > obs
Marshall & van der Lely (2007)	LI	MAE	Complex vs. simple coda	Elicitation	Simple > complex
Leonard et al. (2007)	Both	MAE	Probability	Elicitation	High > low prob (LI only)
Stemberger (2007)	TD	MAE	Probability	LS	High > low prob
Pruitt & Oetting (2009)	TD	AAE	C vs. V	Elicitation	V > C
Blom & Paradis (2013)	Both	MAE/bilingual	Allomorphic	Elicitation	Segmental > syllabic
Riches (2015)	TD	MAE	Sonority slope	Imitation	Falling > level
Owen Van Horne & Green Fager (2015)	Both	MAE	Son vs. obs/alveo	Elicitation	Son > obs/nonalveo > alveo
Third person singular					
Song et al. (2009)	TD	MAE	C vs. V	LS	V > C
Blom et al. (2012)	TD	Bilingual	Allomorphic	LS	[-z] > [-s] > syllabic
Barlow & Pruitt-Lord (2014)	TD	MAE/AAE	C vs. V/son vs. obs	LS	C > V/son > obs (MAE only)

Note. Only significant findings ($p < .05$) for regular forms are listed; LI = group with language impairment; TD = group with typical development; MAE = (mainstream) monolingual American English; Allomorphic = syllabic form versus monosegmental form; Son vs. obs = sonorant versus obstruent; Probability (prob) = phonotactic probability/frequency; LS = language sample; AAE = African American English; C vs. V = consonant versus vowel; Alveo = alveolar versus nonalveolar place of articulation.

[a]Context of higher marking rate > context of lower marking rate.

less frequent across languages or more difficult to acquire or produce) than vowel–consonant sequences (VC or VCVC; e.g., Greenberg, 1978).

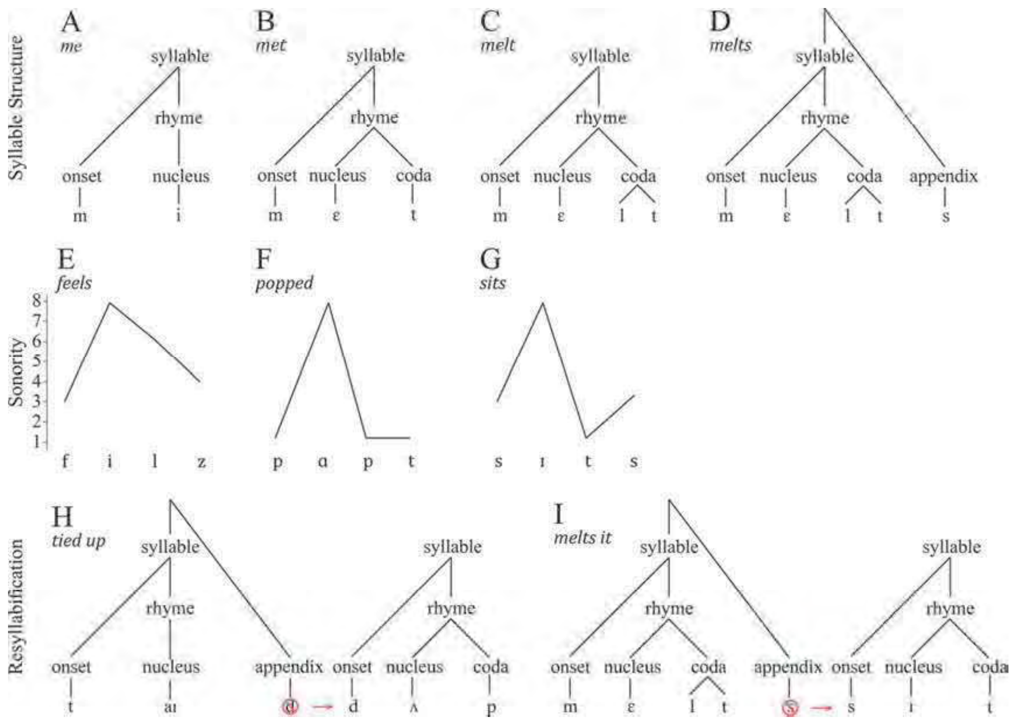
In the case of monosegmental PT or 3s inflection, marking of the morpheme will always result in the addition of a consonant to the word-final sequence. When the verb root ends with a consonant, this creates a word-final consonant cluster (e.g., *step* /stɛp/ → /stɛpt/) and, therefore, a more complex sequence than when the verb root ends with a vowel, which creates a final VC sequence (e.g., *fly* /fla͡i/ → /fla͡iz/). Thus, in terms of sequential complexity, we would expect higher morpheme marking rates when preceded by a vowel (less complex) than when preceded by a consonant (more complex). In fact, PT marking has been found in previous studies to be higher in preceding-vowel contexts (Marshall & van der Lely, 2007; Pruitt & Oetting, 2009).

Structural complexity. Sequential phonological context alone does not fully capture phonological structure. We must also consider the syllable, a hierarchical arrangement that organizes sequential speech segments into naturally parsable groups (e.g., Blevins, 1995; Zec, 2007). Syllable structure is often illustrated using a schematic known as a syllable tree, shown in the first row of Figure 1. For our purposes, it is most important to note that each syllable contains a nucleus, which is usually a vowel, and that most consonants occupy positions outside of the nucleus. Consonants generally occur in the onset if at the front (e.g., /m/ in *met*) and the coda if at the end (e.g., /t/ in *met*) of the syllable. Consonant clusters within any part of the syllable create branches in the syllable structure, which are considered more complex. Thus, in order of increasing complexity, *me* < *met* < *melt,* as shown in Figures 1A–1C.

There are of course exceptions to the prototypical arrangement of syllable structure, with one notable example being word-final consonants that may exist outside of the syllable in the appendix position, as shown in Figure 1D. What is especially relevant for the morphophonologically inclined is that the appendix position has been, at least in part, proposed in explanation of the peculiar behavior of English word-final morphemes within syllables (e.g., Borowsky, 1989; Goldsmith, 2011; Nathan, 2008; Selkirk, 1982). The complexity of the appendix structure is not yet clear. Though three adjacent word-final consonants are sequentially complex, how and if the appendix segment contributes to syllable-structure complexity still merits investigation. Despite these questions, syllable structure is likely an aspect of phonological complexity at play in the formation of words and phrases (e.g., Zec, 2007). In particular, the exceptional nature of word-final morphemes, whatever position they may occupy, suggests that these segments might interact with the preceding syllable structure in ways that might differ from final phonemes that are not morphemes. Thus, the findings for PT by Pruitt and Oetting (2009) and Marshall and van der Lely (2007) can also be attributed to syllable-structure complexity, because a word-final VC sequence denotes a simple coda (as in Figure 1B) and a word-final CC sequence likely requires a complex, branching coda (as in Figure 1C) or an appendix (as in Figure 1D).

Transitional complexity. Phonological complexity also extends to sonority transitions within a syllable. Sonority

Figure 1. Syllable structure (A–D), sonority distance (E–G), and resyllabification (H–I) in contexts of varied word-final complexity.



refers to the relative acoustic energy of a sound segment, often arranged in a hierarchy that ranks each phoneme in a given language from most to least sonorous: vowel > glide > liquid > nasal > voiced fricative > voiceless fricative > voiced stop > voiceless stop (Clements, 1990). Sonority and the difference in sonority rank between adjacent segments have been shown to play an important role in syllable organization and reduction patterns during language acquisition (for a review, see Barlow, 2016). This is due in large part to a cross-linguistic preference for syllables that transition from a low-sonority initial segment to a high-sonority nucleus (an upward sonority slope) and end with a segment that either falls in sonority or remains level (sonority sequencing principle; Clements, 1990). This preferred configuration of sonority slopes is thus considered less complex, whereas sonority transitions that do not adhere to this pattern within a syllable are considered more complex.

The second row in Figure 1 illustrates sonority transitions at the end of a syllable. *Feels* /filz/ demonstrates a falling sonority slope, and *popped* /pɑpt/ presents a level word-final slope. Both demonstrate preferred sonority configurations. Word-final morphemes are known to deviate from the preferred form in that they can form an upward sonority slope when a level or falling slope would be expected according to the sonority sequencing principle, as shown in Figure 1G. In *sits,* the sonority transition from the penultimate segment /t/ to the morpheme segment [-s] creates an upward word-final sonority slope. This violation of syllable-internal transitional properties is also cited as evidence in support of the syllable-external appendix status of English word-final morphemes (Clements, 1990; Fujimura & Lovins, 1978; Halle & Vergnaud, 1980). A number of studies have examined PT marking according to sonority-based influences, consistently identifying marking rates that increase in the context of falling sonority slopes, and therefore phonologically simpler environments (Johnson & Morris, 2007; Oetting & Horohov, 1997; Owen Van Horne & Green Fager, 2015; Riches, 2015).

Probabilities and frequencies. More recent studies have used measures of probability or frequency to describe phonological context, examining the influence of neighborhood density, inflectional frequency, and phonotactic probability on PT marking rate (Leonard, Davis, & Deevy, 2007; Marchman, 1997; Stemberger, 2007). Unlike the aspects of phonological context discussed thus far, these measures all relate to the frequency with which all or part of the tense-inflected verb occurs in speech. In brief, neighborhood density is a measure of the frequency of words that have a similar form, inflectional frequency refers to the number of occurrences of the inflected form of a given verb, and phonotactic probability is derived from the frequency of occurrence of sounds or sequences in a given word position. In these studies examining PT marking, these different measures were used to capture the frequency of occurrence of the inflected form in some fashion. The findings are similar in that they demonstrate higher PT marking rates in more phonotactically probable or frequently inflected contexts. These results on the basis of probability and frequency also suggest that PT is marked more frequently in less phonologically complex environments, given that more frequently occurring or probable contexts might also be considered less complex.

3s Marking and Expansion of Phonological Context

Whereas word-final phonological complexity has shown a relatively consistent pattern of influence on PT marking rates across studies, the few existing studies examining 3s marking and phonological context appear to present conflicting results. Blom et al. (2012) found that bilingual children marked monosegmental [-z] allomorphs more often than [-s], and both more than syllabic [-ɪz] forms—a finding suggestive of the influence of preceding context. In a more direct examination of preceding phonological context in mainstream monolingual English-speaking children, Song et al. (2009) identified higher 3s marking rates in preceding-vowel contexts. On the converse, Barlow and Pruitt-Lord (2014) examined a cross-section from the same corpus used in the longitudinal study of Song et al. and found 3s marking rates to be lower in preceding-vowel contexts. These later findings are unexpected, not only because they differ from a study sampling the same population, but also because they appear contrary to patterns that are expected on the basis of word-final phonological complexity.

Barlow and Pruitt-Lord suggested that differences in control of following context may have contributed to their results that differed from those of Song et al. To date, few studies have examined the effect of the following phonological context on morpheme production (tense or otherwise), even though many have appealed to complexity in syllable structure (e.g., Marshall & van der Lely, 2007; Polite, 2011; Riches, 2015; Song et al., 2009). Excepting utterance-final (UF) productions, we can expect—given resyllabification processes that occur during production of connected speech (Cholin, Schiller, & Levelt, 2004; W. J. M. Levelt, Roelofs, & Meyer, 1999)—that syllable formation will occur across word boundaries in some instances, such as when the preceding word ends with a consonant (as is the case with an inflected regular verb) and the following word begins with a vowel. This tendency, motivated by a universal preference for simple syllables that begin with a consonant and a preference against branching rhymes with a coda (e.g., Barlow & Gierut, 1999), is sometimes referred to as the onset maximization principle (e.g., Selkirk, 1981). For example, it is argued that this principle motivates the occurrence of allomorphy for the *a*/*an* definite article in English. However, this process does not apply equally across languages (Cholin et al., 2004). For instance, this principle is considered highly robust in Spanish (e.g., production of *las alas* /las alas/, "the wings," as [la.salas]), perhaps more so than in English (Colina, 2009; Martínez-Gil, 2000).

The third row in Figure 1 illustrates two examples of following-vowel environments where such resyllabification might occur with PT and 3s morphemes. For simplicity, the PT and 3s morphemes are both presented in the appendix position. Note, however, that evidence for this position is more robust in the presence of an already-complex coda, as in Figure 1I, than in Figure 1H, where the PT morpheme [-d] could also legally occupy a coda position (Selkirk, 1982). Despite disagreement as to whether and how such resyllabification occurs in connected speech[1] (e.g., Jensen, 2000; Szigetvári, 2001), it is reasonable to also consider the following phonological environment, because the morpheme segment may be produced in the onset of the following syllable in some cases.

Some studies have examined the role of following phonological context on morpheme marking rate (plural *-s*: Barlow & Pruitt-Lord, 2014; Polite, 2011; 3s: Barlow & Pruitt-Lord, 2014). These studies failed to identify an effect of following context (consonant vs. vowel) for plural *-s;* however, in comparing 3s and plural *-s* marking rates, the effect of a following vowel was shown to be contrastive. To be specific, 3s marking rate increased when followed by a vowel, yet plural *-s* marking rate decreased in the same context (Barlow & Pruitt-Lord, 2014), suggesting that the influence of following context might differ for 3s and plural *-s*. It is curious that no published studies of following context are available for PT. Nevertheless, given that PT is also likely subject to resyllabification processes, it would be appropriate to expand examination of its phonological context to the following environment as well. In addition, given differences in the influence of following context for 3s and plural *-s* morphemes, examination of following context might reveal differences or similarities between PT and 3s marking patterns.

The Current Study

An overview of the existing research pertaining to the influence of phonological context on tense-marking rates generates questions about differing sensitivity to phonological context across morphemes and populations and the role of resyllabification and following context. In the interest of addressing some of these gaps and questions in an exploratory fashion, we examine phonological influences on both PT and 3s morphemes that include multiple aspects of word-final phonology in naturalistic connected speech samples. We also expand the scope of context to examine phonological influences that might interact with the following word. Further, we contribute to the emerging typology of morphophonological interactions in tense marking by examining these effects in a population of young Spanish–English DLLs.

[1]For instance, Kahn (1976) provides phonetic evidence from the realization of /t/ allophones for the occurrence of ambisyllabic consonants (i.e., those which straddle two syllables) in lieu of resyllabification in some contexts (for a discussion, see Treiman & Zukowski, 1990).

We seek to answer the following questions:

1. Do aspects of phonological context affect the marking rate of PT and 3s morphemes in Spanish–English DLLs with typical development?
2. Does syllable constituency or resyllabification across the word boundary contribute to the influence of phonological context?
3. Which measure or combination of measures of phonological context best predicts morpheme marking rates?
4. Beyond marking rate alone, do the influences of phonological context differ for PT and 3s morpheme productions? Further, are these differences attributable to the phonological form of these morphemes?

On the basis of existing research in other populations, we predict that, across measures, phonological context will affect marking rates for both morphemes. Further, given consistent findings for PT word-final influences, we expect this effect to follow patterns that are based on phonological complexity (i.e., higher marking in less complex contexts), at least for PT. Given inconsistent findings for the 3s morpheme in prior studies, we expect that phonological context will affect PT and 3s morpheme productions differently. And last, we expect resyllabification and following context to affect marking rate, given the notable influence of resyllabification processes in Spanish, the first language of our sampled population.

Method

Participants

Ten Spanish–English bilingual children, aged 56–66 months ($M = 61.7$, $SD = 3.4$) and enrolled in a Southern California public preschool, participated in this study as part of a larger community-based intervention project. All participants accessed the same curriculum in one of two classrooms where a teacher, a teacher's aide, and at least one speech-language pathology student volunteer were present 5 days a week.

Participants were all identified as Spanish–English DLLs with typical development by their caregivers. In written questionnaires, caregivers described their own education level and language use at home. Maternal and paternal education did not exceed a high school diploma. All children were described by caregivers as native speakers of Spanish, and language use at home was further described as percent spoken (output) and percent heard (input) in Spanish and English (e.g., Fabiano-Smith & Goldstein, 2010). At the time of this study, all participants had Spanish input and output at least 30% and 45% of the time at home, respectively (Pearson, Fernandez, Lewedeg, & Oller, 1997). Enrollment in an English-language preschool program denoted consistent English exposure for at least one academic year. Thus, the children in this study present with the profile of a sequential bilingual child (e.g., Hammer et al., 2004), with immersion in their first language, Spanish, from birth and immersion in their second language, English, occurring later in school (starting at age 3–4 years). At this stage, both languages are in development, hence the children's DLL status. However, their second language, English, is in a particularly early stage of development that is highly variable across individuals (Iglesias & Rojas, 2012). Though mastery of PT and 3s morphemes is common in monolingual English learners with typical development by age 4–5 years (Brown, 1973), marking rates for these morphemes in same-age sequential bilingual children are much more variable—ranging anywhere from complete absence to mastery—as attested across studies of young bilingual children (e.g., Bland-Stewart & Fitzgerald, 2001; Gutiérrez-Clellen et al., 2008; Marinis & Chondrogianni, 2010; Padilla, 1978).

Caregivers, with teacher support, also completed the Ages and Stages Questionnaire (Squires & Bricker, 2009) as an indicator of each child's overall development. All scores on this measure exceeded the cutoff for typical development. All children demonstrated nonverbal IQ within the normal range on the Figure Ground and Form Completion subtests of the Leiter International Performance Scale–Revised (Roid & Miller, 1997). In addition, teachers and the lead certified speech-language pathologist working on a community-based intervention in the classrooms identified these children as developing language typically. Thus, in addition to standardized measures of development, typical language development was determined by reports from caregiver, teacher, and speech-language pathologist, which have been shown to provide reliable measures of language development in bilingual populations, for whom norms of many standardized language assessments are inappropriate (Gutiérrez-Clellen & Kreiter, 2003; Paradis, Emmerzael, & Duncan, 2010).

Caregiver questionnaires and nonverbal-intelligence measures were administered at the beginning of the school year, and language samples were collected at the end of the same school year. Standardized testing and language sampling were conducted by speech-language pathology graduate students as part of the larger community-based intervention project.

Procedure

Language-Sample Measures

Spontaneous language samples were the sole source of grammatical-morpheme production measurements in this study. Samples were collected for all children following a play-based elicitation protocol. This protocol included conversation and play with people figures, a gas station, cars, a doll family, toy food, and a picnic set, followed by story retell using three Apricot pictures (Arwood, 1985). Though obligatory 3s and PT contexts occurred throughout the samples, 3s occurred most frequently during the play portion and PT most frequently during story retell. Language samples from 10 participants were orthographically transcribed and analyzed for broad language measures using the Systematic Analysis of Language Transcripts (SALT) software and guidelines (Miller & Iglesias, 2004) as part of the larger community intervention project. Average sample length was 15 min., for a total of 1,993 complete and intelligible utterances (M = 199.3, SD = 84.48). Mean length of utterance in morphemes was 3.61 (SD = 0.41).

Phonological Transcription

A protocol for phonemic transcription and coding of word-final grammatical-morpheme production was developed using Phon (Rose & Hedlund, 2016), a free and open-source software for phonological transcription and analysis. All child-produced utterances were transcribed in their entirety for each language sample in the International Phonetic Alphabet (IPA) using broad phonemic transcription. Target forms were also transcribed to reflect the correct, adultlike form of each utterance. For the purposes of this study, the target form was transcribed to match the child's production except for any errors in the child's production. In other words, the target transcription reflected acceptable idiosyncratic speech variation and dialectal influences, such as flapping of intervocalic alveolar stops, substitution of tense vowels for diphthongs, and dental production of alveolar obstruents.

Utterances were segmented such that any pause in connected speech, as determined by auditory perception and visual interpretation of the waveform and spectrogram, indicated the start of a new utterance. This operationalization of utterance boundaries is phonetic rather than syntactic in nature and is paramount to our analyses of connected speech. For example, *he melts it* /hi mɛlts ɪt/ would be transcribed as a single utterance only if there were no pauses between all three words. Given connected speech (Cholin et al., 2004; W. J. M. Levelt et al., 1999) and onset maximization principles (Clements, 1990; Selkirk, 1981; Zec, 2007), resyllabification could occur across word boundaries in this instance, as shown in Figure 1.

Morpheme Marking

During the transcription process, each utterance was also coded for PT and 3s morphemes in an additional tier within Phon. In this study, only monosegmental regular forms of these morphemes were analyzed (i.e., [-s, -z, -t, -d]), because the VC structure of [-ɪd, -ɪz] allomorphs adds an additional syllable to any word to which they are suffixed. Irregular forms were similarly excluded, because they can vary greatly in terms of which phoneme segments differ between the inflected and bare forms (e.g., *go* → *went*, *be* → *is*).

PT and 3s morphemes were coded at each obligatory context, irrespective of the child's production of the morpheme. In other words, transcriptions were initially coded only for the presence of obligatory contexts of PT and 3s morphemes. From these transcriptions, an obligatory context was considered marked if the morpheme was identifiable in the child's production. Thus, there are contexts in which a child's errant or idiosyncratic production may be coded as marked (see Table 2 for examples). For instance, in one child's production of *he needs to drink,* the inflected verb was produced as [nis]. Though the final consonant of the verb stem was deleted, and the morpheme segment itself was devoiced, the morpheme segment was still phonologically present; thus, this token was coded as marked. In these samples, distortions of the morpheme segment were limited to devoicing or dentalization. On the converse, any obligatory context in which no morpheme segment was identified in the child's production was considered unmarked. Other morpheme-production error types, such as commission or overregularization, were not included in the analyses.

Phonemic transcription and morpheme coding were completed independently of the aforementioned SALT transcriptions by a separate group of graduate and upper-division undergraduate research assistants. Each phonemic transcriber demonstrated consistent and reliable transcription of child speech and identification of word-final morphemes before beginning the project. Transcribers were also trained in the use of Phon, interpretation of spectrograms and waveforms, and allophonic patterns of English and Spanish. Transcribers used both auditory interpretation of the child's speech and visual interpretation of spectrograms and waveforms to guide their transcriptions. Each language sample was transcribed and coded twice by separate transcribers, unaware of each other's coding and transcriptions. Reliability was calculated for 56% of the obligatory marking contexts. Each context included the bare verb stem, the final morpheme segment, and the following word (excepting UF forms). Reliability for IPA transcription of the child productions was 91%; reliability for identification of obligatory 3s and PT contexts was 99%; and reliability for transcription of the PT or 3s morpheme segment (i.e., marked versus unmarked tokens) was 88%. Disagreements were resolved by consensus after both transcriptions were complete. When consensus could not be reached due to noise or poor intelligibility, the item was marked as unintelligible and excluded from the analyses.

Phonological Predictor Variables

In order to examine multiple aspects of phonological context, including those which might extend across the word boundary, we used four different measures of phonological context: one measure of surrounding sequential complexity, two measures specific to preceding (i.e., word-final) phonological context, and one probabilistic measure of morpheme resyllabification into the following word. Using multiple measures of phonological context allowed for better comparison to the results of previous research, given the variability of measures used in those studies (see Table 1), and exploration of the phonological influences beyond preceding context. Furthermore, analyses of multiple measures allowed for comparison of their power to predict morpheme marking.

Number of surrounding consonants. This is the broadest measure, incorporating both preceding and following phonological contexts. This integer variable is a discrete count of the consonants in the coda of the verb stem and the consonants in the onset of the following word. For

Table 2. Contextual examples of past tense and third person singular tokens coded as marked or unmarked.

Orthography	International Phonetic Alphabet (IPA) target	IPA actual	Target → actual	Marked
He needs to go	/hi nidz tu go͡ʊ/	[hi nid tu go͡ʊ]	/z/ → Ø	No
They dropped all the	/ðe͡ɪ dɹɑpt ɑl ðə/	[dʒɚ dɹɑp ɑl ðə]	/d/ → Ø	No
He needs to drink	/hi nidz tu dɹɪŋk/	[hi nis tu dɹɪŋk]	/dz/ → [s]	Yes
Daddy comes to the	/dædi kʌmz tu ðə/	[dædi kʌms tu ðə]	/z/ → [s]	Yes
I fixed it	/a͡ɪ fɪkst ɪt/	[a͡ɪ fɪkst ɪt]	/t/ → [t]	Yes

example, in *jumps by* /dʒʌmps ba͡ɪ/, the final morpheme [-s] is surrounded by three consonants: preceding /m/ and /p/ and following /b/. Because sequences of adjacent consonants are more sequentially complex than fewer consonants or alternating VC sequences, this measure is indicative of surrounding sequential complexity.

Preceding consonant or vowel (C/V) context. Preceding C/V context labels the final segment of the verb stem (i.e., the preceding segment) as a consonant or vowel. For example, in *jumped off* /dʒʌmpt ɑf/, the preceding segment is /p/, and thus the preceding C/V context is C. This measure captures preceding structural complexity in that words with branching coda structures (i.e., multiple word-final consonants) are more complex than words with a simple coda or appendix (i.e., only a single word-final consonant).

Preceding sonority distance. This variable takes into account the difference in sonority level of adjacent segments, also referred to as a sonority slope. Sonority values were coded using the sonority hierarchy for all preceding and morpheme segments to calculate the sonority distance between them. To illustrate, in the example *picked me* /pɪkt mi/, [-t] is the PT morpheme segment and /k/ is the preceding segment. Preceding sonority distance[2] is derived by subtracting the sonority rank (*vowel* [8] > *glide* [7] > *liquid* [6] > *nasal* [5] > *voiced fricative* [4] > *voiceless fricative* [3] > *voiced stop* [2] > *voiceless stop* [1]; Clements, 1990) of the morpheme segment from the preceding segment. In our example, the rank of /t/, a voiceless stop, is subtracted from the rank of /k/, also a voiceless stop ($1 - 1 = 0$). Because there is no change in sonority rank here, the preceding sonority distance is zero. Marking trends identified by this variable reflect transitional complexity, in that a falling word-final sonority slope is simpler than one that remains level or transitions from low to high (Clements, 1990).

Morpheme resyllabification. A final predictor variable addresses the role that resyllabification may play in mediating other phonological influences (i.e., interactions with other predictor variables). This measure is derived from the phonotactic probability of the phoneme sequence that would result if the morpheme segment were to resyllabify into the onset of the following word. Herein, we will refer to this novel measure simply as *morpheme resyllabification*. For this study, all phonotactic-probability measures were calculated using average biphone frequencies from a calculator by Storkel and Hoover (2010), which derives frequencies from a combined corpus of American English speech productions of kindergarten and first-grade children (Kolson, 1960; Moe, Hopkins, & Rush, 1982).

To illustrate, consider the phrase *wants to* /wɑnts tu/. The phonotactic probability of the sequence /stu-/ at the beginning of a word is relatively high because this forms a phonotactically legal sequence, which also occurs at the beginning of many English words. On the other hand, in the phrase *picks the* /pɪks ðə/, the phonotactic probability of the sequence /sðə-/ at the beginning of a word is very low; in fact, it is zero because it does not occur in the onset of any words in the corpus of child speech from which the probability was derived. Further, this sequence is illegal according to English phonotactic constraints (Pitt, 1998; Treiman & Zukowski, 1990). Therefore, contexts where morpheme resyllabification is more probable are presumed to be instances where the morpheme is more likely to be associated with the beginning of the following word, as in *wants to* /wɑnts tu/ but not *picks the* /pɪks ðə/. This is because in these contexts, the morpheme segment could feasibly occupy a position in the following syllable. On the converse, a zero or low morpheme-resyllabification value, as in *picks the*, suggests that the morpheme segment is unlikely to occupy a position in the syllable at the start of the following word because this sequence is either uncommon or nonexistent at the start of words in English.

This measure was selected in consideration of the young DLL population examined in this study. In young children, emerging syllable structure might look very different than in adults (e.g., C. C. Levelt, Schiller, & Levelt, 2000; McLeod, van Doorn, & Reed, 2001), and constraints on syllable structure vary greatly across languages (e.g., Lleó & Prinz, 1996). Not only are the children in this study developing rules for syllable formation, but they are doing so for two languages with differing constraints on syllable structure (e.g., Blevins, 1995). Thus, we used a probability-based measure that does not require overt assumptions about underlying syllable structure or categorical distinctions on the basis of English phonotactics. This measure is derived solely from the relative frequency of occurrence of a resyllabified sequence, as observed in child speech.

[2]This measure would be most accurately described as word-final sonority distance. However, we at times refer to it as preceding sonority distance to differentiate it from other more global measures of context that extend across the word boundary.

Lexical and Sublexical Control Variables

Lexical frequency and phonotactic probability of the bare verb or inflected verb have been shown in previous studies to influence morpheme marking (e.g., Blom et al., 2012; Leonard et al., 2007; Marchman, 1997; Stemberger, 2007). In this study, these properties were not manipulated during data collection, because the language sampling procedure did not include stimulus words; rather, the children were free to produce (or avoid) verbs of their choosing. To address potentially confounding factors related to verb properties, we included word-final phonotactic probability and lexical frequency of the verb stem as control variables.

Word-final phonotactic probability. To control for the phonotactic properties of the final sound sequence of the inflected verb, phonotactic probability was calculated for the vowel of the final syllable, any coda consonants, and the final morpheme (e.g., /ɑps/ in *stops him* /stɑps hɪm/). These phonotactic-probability values were derived using the same corpus and procedures as were used for the morpheme-resyllabification predictor variable.

Verb-stem frequency. This lexical property is a measure of the frequency of occurrence of a word in a given language or population. For this study, lexical frequency of the bare verb stem was collected from the ChildFreq calculator (Bååth, 2010), which draws from approximately 3,500,000 child productions of English words in the North American and UK English corpora of the CHILDES database (for more information on this freely available database of child language, see MacWhinney, 2000). Drawing from the CHILDES database allowed for collection of frequency counts from a large and heterogeneous corpus of English-speaking child productions and, importantly, permitted restriction to an age range matched to the participants in this study.

ChildFreq analyzes lexical frequency in child output only, and thus is not a measure of child-directed speech input. Frequency values from this calculator were derived from the number of occurrences per 1,000,000 words in all English-speaking children between 54 and 71 months of age from the aforementioned CHILDES corpora. Lexical frequency of the verb root (as opposed to the inflected form) was used, because this database did not distinguish between homophonous inflected forms (e.g., 3s *he walks* vs. plural *such long walks*).

Coding Procedures

Marking rate, the dependent variable in this study, is defined as the proportion derived from the number of marked contexts (i.e., where the morpheme segment is present in the child's production) out of the total number of obligatory contexts for a given morpheme. Verb-stem frequency values were collected, unchanged, from the ChildFreq calculator. Phonotactic-probability measures were calculated as the average of the positional biphone frequencies collected from the Storkel and Hoover (2010) calculator. Codes for the remaining variables were generated automatically in Phon or with spreadsheet formulas directly from IPA transcriptions of the target productions for each obligatory PT or 3s morpheme context. Reliability for 25% of the PT and 3s items on the lexical-frequency measure was 100%. Reliability for 25% of the 3s items on two phonotactic-probability measures was 98%. All phonotactic-probability measures for PT were calculated twice because three errors were identified during reliability testing. All errors (four in total) were corrected and confirmed by a third party.

Statistical Analyses

To identify the effect of phonological context on marking rate for two morphemes, examine differences between them should they exist, and identify which phonological influences are most predictive in each case, we compared several families of mixed logistic regression models. A summary of the variables in these analyses is given in Table 3. For each morpheme, a separate family of model comparisons was conducted for each main phonological predictor (number of surrounding consonants, preceding C/V context, and preceding sonority distance). In other words, each main predictor was examined independently. Within each family, morpheme resyllabification was examined as an interacting predictor variable and both word-final phonotactic probability and verb-stem frequency were included as fixed control factors in the initial model. Because morpheme resyllabification is dependent on the following phonological context, all analyses that included this variable excluded UF forms. To account for differences across individuals, participant was included as a random factor in each model. Given repeated observations from each child, standard error was clustered by participant to obtain robust variance estimates (Rogers, 1993). Model variables were not orthogonalized; rather, collinear predictors were examined in a separate model (see Wurm & Fisicaro, 2014).

Because we are interested in identifying the most predictive (or most predictive combination) of these various measures of phonological context, several model comparisons and diagnostic techniques were implemented. Each family of examinations began with the most complex model, including both control factors and the interaction between the main phonological predictor for that family and morpheme resyllabification. Interactions and control factors that were not significant ($p \geq 0.05$), did not contribute to goodness of fit as determined by Akaike's information criteria (Akaike, 1974; Sakamoto, Ishiguro, & Kitagawa,

Table 3. Regression variables.

Predictor
Number of Surrounding Consonants × Morpheme Resyllabification
Preceding C/V Context × Morpheme Resyllabification
Preceding Sonority Distance × Morpheme Resyllabification
Control
Word-final phonotactic probability
Verb-stem frequency
Participant (random factor)

1986), and/or led to poor specification (as determined by a model-specification link test[3]) were removed. The results reported for each family of comparisons therefore reflect the most parsimonious model, and thus the most predictive and nonredundant configuration available for each measure of phonological context.

Results

Marking of PT and 3s morphemes was sufficiently variable for observation of phonological influences on marking rates, though data sampling from spontaneous connected speech resulted in nearly twice the number of obligatory contexts for 3s ($n = 92$) than for PT ($n = 54$). Average marking rate for PT was 27.8% ($SD = 45.2\%$), with large variation across participants ($SD = 33.1\%$). Marking rate for 3s was higher at 47.8% ($SD = 50.0\%$), also with considerable variation across participants ($SD = 29.0\%$). These findings are congruent with highly variable word-final morpheme marking rates in other studies of Spanish–English DLLs (Bland-Stewart & Fitzgerald, 2001; Gutiérrez-Clellen et al., 2008; Marinis & Chondrogianni, 2010; Padilla, 1978). Four participants did not produce PT in any obligatory contexts, though they evidenced multiple productions in PT probes and other standardized language measures, demonstrating some knowledge of the morpheme and ability to produce it; thus, unmarked contexts for these participants were included in the PT analyses. Two participants did not produce 3s in any obligatory contexts, and no data were available to otherwise demonstrate their ability to produce the morpheme; thus these participants were excluded from the 3s analyses. Results for each family of comparisons, for each morpheme, are displayed in Figure 2.

Number of Surrounding Consonants

There was a significant negative effect of number of surrounding consonants on PT marking rate, $z(1, 54) = -2.47$, $p = .01$, and no interaction with morpheme resyllabification. PT marking was at 100% (of five instances) in the simplest phonological context of no surrounding consonants (preceded by a vowel and in UF position). Note that a PT morpheme surrounded by a preceding and a following vowel (V_V) would also constitute a context with no surrounding consonants; however, this context did not occur in these samples. Marking rates for more complex contexts of one or two surrounding consonants were much lower, and all (four) instances of PT surrounded by three consonants were unmarked. Though the most parsimonious model available[4] was used, a significant specification link test (see Table 4) suggests that variables in this model may have been poorly specified. This could be related to few vowel-final PT verb stems and the absence of V_V context in these samples or the influence of UF position.[5] Thus, this result should be interpreted with caution because the predictive power of this measure might be affected by sampling limitations or unidentified context effects.

Overall, number of surrounding consonants had a significant positive effect on 3s marking rate, $z(1, 78) = 4.73$, $p < .01$, and this influence was modulated by the probability of morpheme resyllabification—that is, a significant Number of Surrounding Consonants × Morpheme Resyllabification interaction, $z(1, 78) = -3.46$, $p < .01$—as determined by the most parsimonious model.[6] The simple main effect of morpheme resyllabification was also significant, $z(1, 78) = 2.57$, $p = .01$.

Examination of the interaction between surrounding consonants and morpheme resyllabification reveals that the total number of surrounding consonants does influence 3s marking rate, but this influence depends on the probability of morpheme resyllabification into the following word. In contexts with low morpheme resyllabification, a greater number of surrounding consonants results in higher 3s marking rates (an unexpected pattern on the basis of sequential complexity). In high-resyllabification contexts, increasing surrounding consonants results in lower 3s marking (the expected pattern also observed for PT).

Preceding C/V Context

Preceding C/V context significantly influenced PT marking rate, $z(1, 54) = 5.62$, $p = .02$, and did not interact with morpheme resyllabification, as determined by the most parsimonious model.[7] PT marking rate increased when preceded by a vowel (i.e., in a simple coda) and decreased when preceded by a consonant.

Preceding C/V context also significantly influenced 3s marking rate, $z(1, 78) = 5.88$, $p = .02$, and, once again, this effect was mediated by morpheme resyllabification, $z(1, 78) = 11.52$, $p < .01$, as determined by the most parsimonious model.[8] The simple main effect of morpheme resyllabification

---

[3]A model-specification link test is used to confirm that a given model is both predictive and unlikely to exclude other significant predictors, interactions, or polynomial transformations. This test generates a model with only two variables: one from the predictions of the model we are examining, and a second that is an exponential transformation of the predictions. Should the exponential transformation reach a level of significance, this would indicate misspecification, such as missing variables and interactions or incorrect formatting of existing variables (Pregibon, 1980).

[4]Independent predictor: number of surrounding consonants. Controls: verb-stem frequency, word-final phonotactic probability. Wald $\chi^2(3, 54) = 7.64$, $p = .05$.

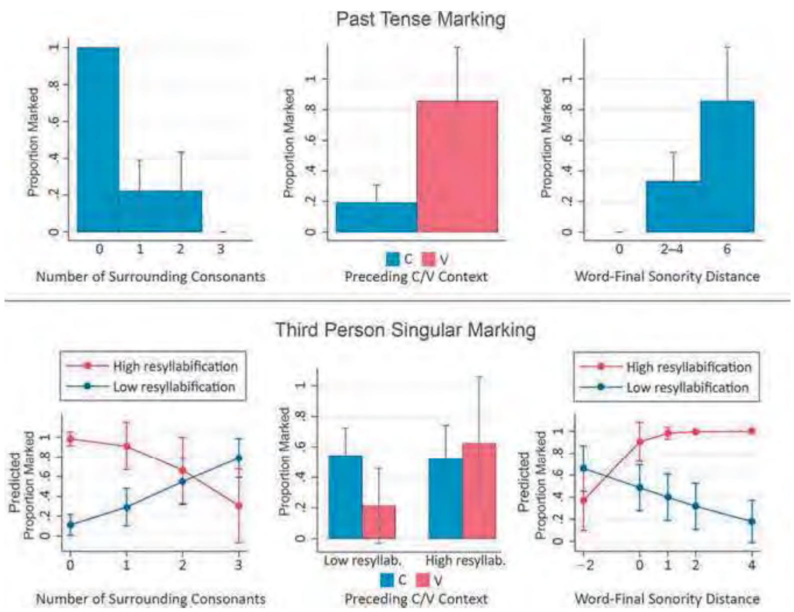
[5]Two follow-up examinations tested this model with UF items excluded and, alternatively, with UF position included as a predictor variable. Exclusion reduced specification error ($p = .60$), though UF position was not a significant predictor ($p = .96$).

[6]Interacting predictors: Number of Surrounding Consonants × Morpheme Resyllabification. Control: verb-stem frequency. Wald $\chi^2(4, 78) = 45.03$, $p < .01$.

[7]Independent predictor: preceding C/V context. Controls: verb-stem frequency, word-final phonotactic probability. Wald $\chi^2(3, 54) = 12.65$, $p = .01$.

[8]Interacting predictors: Preceding C/V Context × Morpheme Resyllabification. Controls: verb-stem frequency, word-final phonotactic probability. Wald $\chi^2(6, 78) = 78.30$, $p < .01$.

Figure 2. Findings for marking rates of past tense and third person singular. For number of surrounding consonants and word-final sonority distance, low resyllabification = morpheme resyllabification at 0.00, and high resyllabification = morpheme resyllabification at 0.01. For preceding consonant/vowel (C/V) context, Low resyllab. and High resyllab. groups represent a median split of the morpheme-resyllabification values (Low resyllab. ≤ 0.0005, High resyllab. > 0.0005). Error bars indicate 95% confidence intervals.



was not significant, $z(1, 78) = -1.77$, $p = .08$. Overall 3s marking rate was higher when preceded by a consonant than by a vowel. However, this effect differed according to probability of morpheme resyllabification. In contexts with low probability of morpheme resyllabification, 3s marking increased with a preceding consonant. In contexts with high probability of resyllabification, 3s marking rate favored a preceding vowel.

Preceding Sonority Distance

As with number of surrounding consonants and preceding C/V context, preceding sonority distance also significantly influenced PT marking rate, $z(1, 54) = 3.01$, $p < .01$. Unlike with previous families of observations, no control factors contributed to the most parsimonious model.[9] When the PT morpheme created larger word-final sonority slopes, marking rate increased. Out of 20 obligatory contexts for PT marking where the verb stem ended in a stop consonant (e.g., *stop[-t]*), thus creating a level sonority slope, it is notable that no participant produced the morpheme segment. In contexts with a slight-to-moderate sonority slope of 2 to 4 (verbs ending with a fricative or nasal; e.g., *clean[-d]*), the PT morpheme was marked in 33% of obligatory contexts. When preceded by a vowel (e.g., *tie[-d]*), which creates a steep sonority slope, PT marking rate was 86%.

Preceding sonority distance also significantly affected 3s marking rate, $z(1, 78) = -2.03$, $p = .04$, and this influence interacted with morpheme resyllabification, $z(1, 78) = 3.86$, $p < .01$, as determined by the most parsimonious model.[10] The simple main effect of morpheme resyllabification was also significant, $z(1, 78) = -2.61$, $p = .01$. Thus, word-final sonority transitions affect 3s marking rate, and this influence is dependent on the probability of resyllabification into the following word.

As with each phonological predictor of 3s marking, two opposing patterns emerged. When resyllabification into the following word was less probable, more steeply falling (and thus, preferred) sonority slopes reduced 3s marking rate. In other words, in these environments the influence of word-final sonority distance is contrary to expectations that are based on transitional phonological complexity. When resyllabification into the following word is more probable, however, more steeply falling sonority slopes increased 3s marking rate, which is the expected pattern on the basis of transitional phonological complexity. This is also, incidentally, the pattern observed for PT.

Phonological Predictors of Tense-Morpheme Marking

Number of surrounding consonants, preceding C/V context, and preceding sonority distance each emerged as significant predictors of tense-morpheme marking rate. In order to identify which measurement of phonological

[9] Wald $\chi^2(1, 54) = 9.08$, $p < .01$.

[10] Interacting predictors: Preceding Sonority Distance × Morpheme Resyllabification. Controls: verb-stem frequency, word-final phonotactic probability. Wald $\chi^2(5, 78) = 22.29$, $p < .01$.

Table 4. Model comparisons for measures of phonological context.

Phonological context	AIC[a]	BIC[a]	LT pred.[b]	LT error[c]
Past tense				
Surrounding consonants	64.88	74.82	> .01	.01[d]
Preceding C/V	60.05	69.99	> .01	.63
Preceding sonority distance	**51.49**	**57.46**	**> .01**	**.74**
Third person singular				
Surrounding Consonants × Morpheme Resyllabification	**98.70**	**112.84**	**> .01**	**.74**
Preceding C/V × Morpheme Resyllabification	100.03	116.53	> .01	.52
Preceding Sonority Distance × Morpheme Resyllabification	102.47	118.97	> .01	.80

Note. Best-fit models are in boldface; AIC = Akaike's information criterion; BIC = Bayesian information criterion; LT = link test; C/V = consonant/vowel.

[a]Lower value = better fit. [b]Link test: significance of model prediction ($p \leq .05$ = significant model prediction). [c]Link test: specification error ($p \leq .05$ = significant specification error). [d]Significant model-specification link test indicating specification error.

context is most predictive for each morpheme, we repeated the same model comparison and diagnostic techniques used in each family of examinations for the most parsimonious model from each set. The results of these comparisons for each morpheme are listed in Table 4. For PT, word-final sonority distance—alone—emerged as the most salient phonological predictor of PT marking rate. For 3s, total number of surrounding consonants and its interaction with probability of morpheme resyllabification were the most predictive aspects of phonological context.

Differentiating Patterns of Tense-Morpheme Marking

Given differences in PT and 3s marking patterns across phonological contexts, an additional analysis was conducted to determine if these differences might be attributable to phonological features of the morphemes themselves. For this comparison, we examined nonmorphemic word-final homophones of both morphemes (e.g., the /d/ in *slide* as compared to the PT [-d] in *cried*; the /s/ in *axe* as compared to the 3s [-s] in *packs*). A randomly selected subset of 78 nonmorphemic final /s/ and /z/ fricatives and 54 nonmorphemic final /d/ and /t/ stops were extracted from the participants' language samples[11] and analyzed for the influence of preceding C/V context following the same procedures as for the word-final morphemes. No frequency or probability data were collected for nonmorphemic segments, and variables derived from these values (morpheme resyllabification and both control factors) were not used in this analysis.

The effect of preceding C/V context on nonmorphemic final-consonant marking rate is compared to the effects observed for PT and 3s in Figure 3. Preceding C/V context did not have a significant influence on the production of nonmorphemic final consonants /s/, /z/, /t/, and /d/, regardless of type, $z(1, 132) = 0.80$, $p = .43$. Similar marking patterns for nonmorphemic final consonants (i.e., /t/ and /d/ do not pattern differently from /s/ and /z/), $z(1, 132) = 0.85$, $p = .40$, suggest that the observed differences in PT and 3s marking patterns cannot be explained by differences between stop (/t/ and /d/) and fricative (/s/ and /z/) sound types.

Discussion

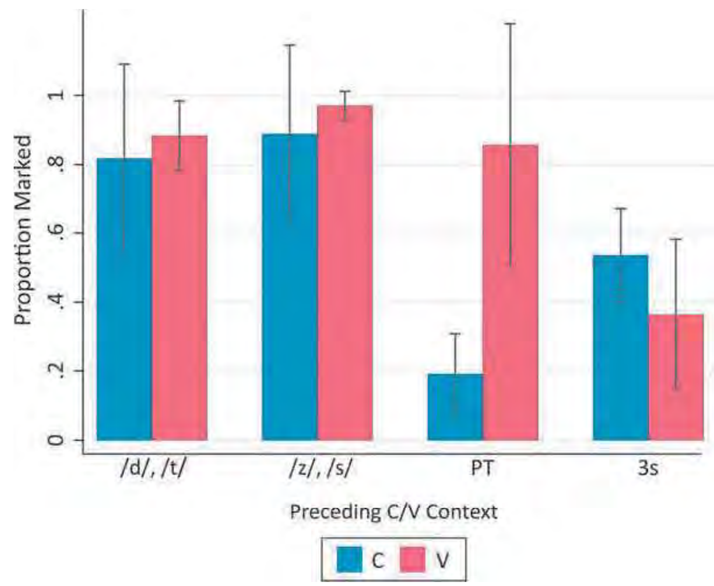
In this study, the influence of phonological context on tense-morpheme marking was examined in a group of preschool-age Spanish–English DLLs. Results are summarized in Table 5. For both PT and 3s morphemes, phonological influence was most strongly identified at the end of the verb stem to which the morpheme attaches. This phonological influence was confirmed by both C/V distinction and word-final sonority configuration—a reassuring finding because it aligns with, and therefore supports, word-final context being the focus for most prior research. We thus extend evidence demonstrating the influence of preceding phonological context on tense-morpheme marking rate to a population of young Spanish–English DLLs.

In addition to the influence of preceding context, the influences of following and surrounding context were also considered, and patterns were compared across PT and 3s morphemes. Upon comparison, PT and 3s morphemes diverge tremendously in the marking patterns that emerged on the basis of their sensitivity to contextual phonological influences. The marking pattern observed for PT supports previous findings in populations of monolingual English speakers. The rate of PT marking increased when preceded by a vowel, when surrounded by fewer consonants, and when forming a steeper word-final sonority slope. Other than the influence of total surrounding consonants, following phonological context was not found to significantly affect PT marking rate. Of all phonological measures examined, word-final sonority slope was the most powerful predictor of PT marking rate.

Examination of phonologically influenced marking patterns for the 3s morpheme yielded notably different results. When considering only preceding phonological context, 3s

[11]Items were sampled to match the frequency of analyzed PT and 3s morphemes. Because all 3s analyses included a significant interaction with morpheme resyllabification (a measure that does not apply to UF forms), UF forms were also excluded from the nonmorphemic homophone analysis of /z/ and /s/.

Figure 3. Marking rate by preceding consonant/vowel (C/V) context for past tense (PT) and third person singular (3s) morphemes and their corresponding nonmorphemic homophones (/d/, /t/; /z/, /s/) in word-final position. Error bars indicate 95% confidence intervals.



marking rate appeared to decrease in the same contexts where PT marking rate was shown to increase. That is, 3s overall marking rate was reduced in the context of preceding vowels, fewer surrounding consonants, and when forming steeper word-final sonority slopes. These findings are compatible with those reported by Barlow and Pruitt-Lord (2014) for younger monolingual English-speaking children but appear in contrast to the findings of Song et al. (2009), where 3s marking rate was shown to increase in the context of a preceding vowel. However, these different findings may be explained by the role of following and surrounding context, which significantly affected 3s marking rate in the

Table 5. Summary of findings for phonological influences on marking rates for past tense and third person singular.

Measure of preceding or surrounding context	Morpheme resyllabification[a]	Simple	Complex
Past tense			
Surrounding consonants	High	✓	
	Low	✓	
Preceding C/V context	High	✓	
	Low	✓	
Preceding sonority distance	High	✓	
	Low	✓	
Third person singular			
Surrounding consonants	High	✓	
	Low		✓
Preceding C/V context	High	✓	
	Low		✓
Preceding sonority distance	High	✓	
	Low		✓

Note. Check mark indicates context (in terms of relative phonological complexity) with the highest marking rate. Simplest contexts for each measure (in descending order): zero surrounding consonants, preceding vowel, and large word-final sonority distance; C/V = consonant/vowel.

[a]Morpheme resyllabification indicates phonotactic probability of resyllabification into the following word. This variable did not interact with other measures for past tense; there were significant interactions with each predictor variable for third person singular ($p < .05$).

current study. Marking of 3s was sensitive to the number of surrounding consonants and syllable-level phonotactic effects across the word boundary. In fact, the most powerful predictor of 3s marking rate was number of surrounding consonants, but only when the probability of resyllabification into the following word was also taken into account. To be specific, two patterns emerged in juxtaposition. When resyllabification of the 3s morpheme into the following syllable was improbable (as determined by phonotactic probability), marking rate increased in complex phonological contexts; in contrast, for PT the marking rate was consistently lower in complex contexts. When resyllabification was more probable, however, 3s marking rate favored simpler phonological contexts in much the same way as PT-marking trends.

The emergence of two distinct patterns of 3s marking in the same preceding contexts suggests that examination of preceding context alone might result in variable patterns across studies if following phonological context (specifically, syllable-level phonotactics) is not taken into account. This could contribute to the different findings for preceding context by Barlow and Pruitt-Lord and by Song et al. In fact, because these studies differed in how they restricted following context, perhaps one finding is not contrary to the other. It may be that they reveal complementary patterns of influence due to sampling different surrounding contexts.

The PT and 3s morphemes have many similarities, such as analogous allomorphic alternations, similar periods of acquisition (in monolingual individuals: Brown, 1973; in Spanish–English bilingual individuals: Bland-Stewart & Fitzgerald, 2001), and Level 2 morpheme status (i.e., monosegmental morphemes that suffix at the word level; e.g., Goldsmith, 2011), and both are tense markers used as diagnostic indicators of LI. These similarities make the observed dissimilar influence of phonological context on their marking rates all the more unexpected. Because the monosegmental forms examined in this study are composed of different sets of phonemes, each with different phonological properties, phonological influences on each morpheme were also compared with influences on their nonmorphemic word-final homophone counterparts. All four nonmorphemic word-final consonants showed the same pattern in favor of simpler phonological contexts. Thus, the differences observed for PT and 3s cannot be attributed only to differences in the types of phonemes these morphemes contain.

Modulating Word-Final Complexity

In this study, word-final phonological context was examined using multiple measures, many of which have appeared in previous studies (see Table 1). Across measures, two types of patterns emerged that can be characterized by preceding and surrounding phonological complexity. For PT and for 3s (though only in the context of high morpheme resyllabification), marking rate increased in phonologically simpler environments. A second, contrastive pattern emerged, but only for 3s in contexts of low morpheme resyllabification. In these environments, 3s not only appeared resistant to the negative influence of phonological complexity but in fact was more frequently marked.

It is argued that word-level morphemes, such as PT and 3s, can occupy an extrasyllabic appendix position, particularly in more complex preceding contexts (Selkirk, 1982). A consideration of extrasyllabic constituency can provide an explanation of these disparate 3s marking patterns. When the morpheme segment is situated outside of the syllable (i.e., appendix position), syllable-internal complexity constraints may not apply to it (Clements, 1990; Selkirk, 1982), effectively fortifying the morpheme against the inhibitory influence of preceding complexity. If the extrasyllabic appendix position fortifies marking in complex contexts, it stands to reason that resyllabification out of this position into the following word would have an impact on this effect.

On the other hand, in simple preceding contexts there is less linguistic motivation for the morpheme to occupy an appendix position, if only because the morpheme segment is unlikely to violate word-final phonotactic constraints in a simple coda (e.g., Selkirk, 1982; Zec, 2007). Therefore, we might not expect the same fortifying effect of extrasyllabic constituency in simpler contexts. Indeed, when neither option results in appendix occupancy, preference (in the form of higher marking rate) is given to the morpheme in onset position over coda position. This is not surprising, given a widely attested cross-linguistic preference for syllables with a consonant onset and a preference against syllable codas (e.g., McCarthy & Prince, 1995). Thus, marking patterns related to the occupancy or vacancy of an appendix position might be involved in the findings for 3s presented here. It is interesting that like 3s, the PT morpheme is also allowed to occur in an extrasyllabic appendix position, though it could be argued that evidence for 3s appendix occupancy is more robust than for PT, given that only 3s can create negative word-final sonority slopes that violate the sonority sequencing principle. Related or not, no such influence of morpheme resyllabification was observed for PT. This description on the basis of appendix constituency is thus only extended to 3s, but a more substantial explanation related to syllabic constituency would need to accommodate observations for both morphemes.

Limitations and Future Directions

In this exploratory study, data were drawn from spontaneous connected speech to minimize a task-related influence and observe naturalistic connected speech patterns. This sampling method, however, was a source of methodological limitation due to the small sample size and the lack of controlled stimulus items. Given the sample size, there may be influences that did not reach significance but still contributed to marking rate, particularly for PT. Further, many potentially illuminating analyses of phonological context were not conducted due to the lack of multiple target productions in each desired condition. The entire surrounding phonological context was suspected to influence marking rates, and indeed, this was confirmed in the form of a morpheme-resyllabification measure for 3s. However, measures of following context that paralleled measures of preceding context (e.g., following C/V context or following sonority distance) were not analyzed because identification of an interaction between preceding and following contexts would require more observations than were available in these data. For instance, preceding and following sonority distance each contain up to eight levels, and a complete examination of their interaction would require observations in as many as 64 conditions. This limitation is particularly true for PT, where vowel-final regular forms (e.g., *cried*; $n = 7$) were especially sparse, and even with evenly distributed UF productions and following consonants and vowels, there would be a maximum of only one or two observations in each combination of C/V context (V_V, V_C, C_UF, etc.).

For this same reason, a probability-based measure of following context (i.e., morpheme resyllabification) was especially useful for detecting influences that transcend the limitations of somewhat arbitrary categorical linguistic distinctions, such as consonant versus vowel. This continuous variable allowed for trend-like interactions that did not require observations in each combination of levels, as is the case for categorical interactions. Nonetheless, limited observations severely restricted analyses of surrounding context. It is possible, if not likely, that following context exhibits a demonstrable impact on tense-marking rate that could be identified with linguistic measures other than morpheme resyllabification.

Elicitation or imitation probes might allow for better control and inclusion of multiple observations in each condition. However, children have been shown to perform differently in probes and spontaneous language samples (Oetting et al., 2012). Moreover, prior identification of a task effect that interacted with the influence of preceding phonological context on PT marking (Barlow et al., 2015) suggests that this difference in performance may also extend to the influence of phonological environment, complicating the use of probes in research of this nature. Future research might systematically examine the effect of task type on morphological sensitivity to phonological context in order to guide data-collection methods in future morphophonological studies.

The influence of a probabilistic measure of resyllabification into the following onset adds to the growing body of research suggesting that lexical and sublexical frequency and probability measures do affect morpheme production (e.g., Hoover & Storkel, 2013; Leonard et al., 2007; Marchman, 1997; Stemberger, 2007). Further, it extends this effect to the word-final morpheme in combination with the initial sequence of the following word. Nonetheless, this measure of resyllabification probability is novel, and thus has yet to be validated. It is important that in a given production, such as *picks it* [pɪks.ɪt], the high phonotactic probability of an onset sequence /sɪ-/, derived from the morpheme-resyllabification measure, may suggest that resyllabification is plausible (as in *pik sit* [pɪk.sɪt]), but it certainly does not indicate if such a syllable formation did in fact occur in that particular production. This measure is essentially a calculation of phonotactic probability, and its extension to online syllable formation is currently only suggestive. Future research could identify the relationship between such a probability measure and syllable formation, for example, in a productive syllable-manipulation task.

This study draws from the connected speech productions of preschool-age Spanish–English DLLs. This is useful in that it expands our typological understanding of morphophonological interactions; however, it limits comparison of these findings to other studies until further research is conducted on young bilingual and monolingual comparison groups. It is not yet clear if these results are particular to the sampled population of Spanish–English bilingual children or generalizable to other bilingual or monolingual populations.

There is reason to consider the linguistic profile of this population, because dual-language interaction in these young learners could ostensibly affect PT and 3s marking patterns differently, especially across word boundaries. Nonpeninsular dialects of Spanish, for example, do not contain word-final grammatical morphemes similar in form to PT /-d/, but they do contain such morphemes similar to 3s /-z/, such as plural /-s/ (e.g., *casas,* “houses”) and present tense second person singular /-es/, /-as/, or /-is/ (e.g., *corres* /kores/, “(you) run”). If Spanish morphology has any influence on these children’s English productions, it might then be expected to surface in similar /s/-type morphemes, such as 3s, but not PT.

Spanish syllables are also highly restricted in their allowance of syllable codas, with an especially strong preference for open CV syllables, both of which contribute to robust resyllabification across word boundaries in Spanish (e.g., Colina, 2009; Martínez-Gil, 2000). In addition, /s/ is subject to Spanish-specific phonotactic constraints, such as impermissibility in onset clusters, which further govern its behavior during resyllabification (e.g., Lipski, 1999). It might then be that resyllabification influences identified for English 3s but not PT are related to particularly stringent constraints against /s/-initial clusters in Spanish. Indeed, Spanish–English bilingual individuals have been shown to activate Spanish-specific phonotactic constraints even when accessing English words (Freeman, Blumenfeld, & Marian, 2016). Future research might isolate and compare the influence of following context on the morpheme productions of Spanish–English bilingual speakers against the productions of English monolingual speakers.

Different findings for PT and 3s sensitivity to phonological context could be attributable to a number of factors and are likely a combination of many. Perhaps this cross-sectional examination is revealing each morpheme in a different state of emergence (note that overall marking rate was higher for 3s than PT in this sample), or perhaps there is some categorical difference in the underlying representation of these morphemes that has yet to be identified. If these morphemes are indeed sensitive to different phonological influences, this knowledge might be used to develop more sensitive diagnostic measures for identification of LI using tense-marking patterns (rather than marking rate alone), especially in linguistically diverse populations. Given that both morphemes were more frequently marked in some phonological contexts and less marked in others, there is reason to further consider the role of phonology in the development and interpretation of diagnostic tools that examine morpheme production. For instance, it may be worthwhile to provide opportunities for word-final morpheme production in a variety of surrounding phonological contexts to provide a more complete demonstration of the child's language use patterns.

Many studies have demonstrated the efficacy of treating complex word-initial sequences for speech sound disorders (e.g., Gierut, 2007). However, there is a poverty of such research for word-final sequences, which is perhaps more outstanding given the frequent co-occurrence of speech sound disorders and LI (e.g., Shriberg & Austin, 1998). Given the sensitivity of word-final tense marking to phonological environment, treatment of word-final sequences could be a promising avenue for future research related to efficacious and even simultaneous treatment of speech sound production and morphosyntax. The findings presented here suggest that multiple phonological factors, including aspects of the following word, influence morpheme marking in a bilingual population. The presence of these influences may have implications for the development and administration of diagnostic tools for identification of LI and for morphophonological treatment approaches, such that future research should consider a broader scope of phonological influence in other linguistically diverse populations, including children with LI.

To conclude, aspects of phonological context and complexity surrounding PT and 3s tense morphemes were examined for their influence on marking rate in young sequential Spanish–English bilingual children. Marking of both morphemes was sensitive to word-final phonological context, though they differed in the nature of this sensitivity. Marking of 3s in particular was also influenced by aspects of following phonological context. We suggest that resyllabification and positioning of the morpheme in syllable structure may contribute to different sensitivities to phonological context. Further research is needed to understand the differences observed for PT and 3s morphemes and the role of a dual-language profile in sensitivity to phonological influences on morpheme production.

Acknowledgments

Philip N. Combiths and Irina Potapova were supported by National Institutes of Health Training Grant T32 DC007361. Additional funding was provided by National Institute on Deafness and Other Communication Disorders RO3 Grant DC012141 and a Price Philanthropies research grant, awarded to Sonja Pruitt-Lord.

The authors would like to thank our young participants and research assistants in the Phonological Typologies Lab and the Child Language Development, Disorders, and Disparities Lab at San Diego State University for their integral role in the completion of this project. We thank Jörg Matt for statistical consultation and Yvan Rose and Greg Hedlund for technical assistance and their role in the development of Phon, as well as attendees of the 2016 International Child Phonology Conference and the 2016 Symposium on Research in Child Language Disorders for valuable feedback on preliminary data. We also thank Sam Engel for contributions to the initial stages of this investigation.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

Arwood, E. (1985). *Apricot I Kit*. Tigard, OR: Apricot.

Bååth, R. (2010). ChildFreq: An online tool to explore word frequencies in child language [Technical report]. *LUCS Minor, 16,* 1–6.

Barlow, J. (2016). Sonority in acquisition: A review. In M. J. Ball & N. Müller (Eds.), *Challenging sonority: Cross-linguistic evidence* (pp. 295–336). Sheffield, United Kingdom: Equinox.

Barlow, J. A., & Enríquez, M. (2007). Theoretical perspectives on speech sound disorders in bilingual children. *Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse Populations, 14*(2), 3–10. https://doi.org/10.1044/cds14.2.3

Barlow, J. A., & Gierut, J. A. (1999). Optimality theory in phonological acquisition. *Journal of Speech, Language, and Hearing Research, 42,* 1482–1498.

Barlow, J. A., & Pruitt-Lord, S. (2014). The role of phonological context in children's overt marking of '-s' in two dialects of American English. In A. W. Farris-Trimble & J. A. Barlow (Eds.), *Perspectives on phonological theory and acquisition: In honor of Daniel A. Dinnsen* (pp. 133–154). Philadelphia, PA: Benjamins.

Barlow, J. A., Pruitt-Lord, S. L., & Combiths, P. (2015, June). *Context and task effects on English past-tense marking: The case of bilinguals*. Paper presented at the International Child Phonology Conference, St. John's, Newfoundland, Canada.

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41,* 1185–1192.

Berko, J. (1958). The child's learning of English morphology. *Word, 14,* 150–177. https://doi.org/10.1080/00437956.1958.11659661

Bland-Stewart, L. M., & Fitzgerald, S. M. (2001). Use of Brown's 14 grammatical morphemes by bilingual Hispanic preschoolers: A pilot study. *Communication Disorders Quarterly, 22,* 171–186. https://doi.org/10.1177/152574010102200403

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

Blom, E., & Paradis, J. (2013). Past tense production by English second language learners with and without language impairment. *Journal of Speech, Language, and Hearing Research, 56,* 281–294. https://doi.org/10.1044/1092-4388(2012/11-0112)

Blom, E., Paradis, J., & Duncan, T. S. (2012). Effects of input properties, vocabulary size, and L1 on the development of third person singular *-s* in child L2 English. *Language Learning, 62,* 965–994. https://doi.org/10.1111/j.1467-9922.2012.00715.x

Borowsky, T. (1989). Structure preservation and the syllable coda in English. *Natural Language & Linguistic Theory, 7,* 145–166.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Cholin, J., Schiller, N. O., & Levelt, W. J. M. (2004). The preparation of syllables in speech production. *Journal of Memory and Language, 50,* 47–61. https://doi.org/10.1016/j.jml.2003.08.003

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, United Kingdom: Cambridge University Press.
Colina, S. (2009). *Spanish phonology: A syllabic perspective*. Washington, DC: Georgetown University Press.
Dalal, R. H., & Loeb, D. F. (2005). Imitative production of regular past tense -ed by English-speaking children with specific language impairment. *International Journal of Language & Communication Disorders, 40,* 67–82. https://doi.org/10.1080/13682820410001734163
Fabiano-Smith, L., & Goldstein, B. A. (2010). Phonological acquisition in bilingual Spanish–English speaking children. *Journal of Speech, Language, and Hearing Research, 53,* 160–178. https://doi.org/10.1044/1092-4388(2009/07-0064)
Freeman, M. R., Blumenfeld, H. K., & Marian, V. (2016). Phonotactic constraints are activated across languages in bilinguals. *Frontiers in Psychology, 7,* 702. https://doi.org/10.3389/fpsyg.2016.00702
Fujimura, O., & Lovins, J. (1978). Syllables as concatenative phonetic units. In A. Bell & J. B. Hooper (Eds.), *Syllables and segments* (pp. 107–120). New York, NY: North-Holland.
Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16,* 6–17.
Goldsmith, J. (2011). The syllable. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (2nd ed., pp. 164–196). Oxford, United Kingdom: Wiley-Blackwell.
Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant clusters. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language, Vol. 2: Phonology* (pp. 243–279). Stanford, CA: Stanford University Press.
Gutiérrez-Clellen, V. F., & Kreiter, J. (2003). Understanding child bilingual acquisition using parent and teacher reports. *Applied Psycholinguistics, 24,* 267–288. https://doi.org/10.1017/S0142716403000158
Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29,* 3–19. https://doi.org/10.1017/S0142716408080016
Halle, M., & Vergnaud, J.-R. (1980). Three dimensional phonology. *Journal of Linguistic Research, 1,* 83–105.
Hammer, C. S., Miccio, A. W., & Rodriguez, B. L. (2004). Bilingual language acquisition and the child socialization process. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (pp. 21–50). Baltimore, MD: Brookes.
Hawkins, J. A. (1987). Implicational universals as predictors of language acquisition. *Linguistics, 25,* 453–473.
Hoover, J. R., & Storkel, H. L. (2013). Grammatical treatment and specific language impairment: Neighbourhood density & third person singular -*s*. *Clinical Linguistics & Phonetics, 27,* 661–680. https://doi.org/10.3109/02699206.2013.789928
Iglesias, A., & Rojas, R. (2012). Bilingual language development of English language learners: Modeling the growth of two languages. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (2nd ed., pp. 1–30). Baltimore, MD: Brookes.
Jacobson, P. F., & Schwartz, R. G. (2005). English past tense use in bilingual children with language impairment. *American Journal of Speech-Language Pathology, 14,* 313–323.
Jakobson, R. (1968). *Child language, aphasia and phonological universals* (A. R. Keiler, Trans.). The Hague, the Netherlands: Mouton.
Jensen, J. T. (2000). Against ambisyllabicity. *Phonology, 17,* 187–235.
Johnson, B. W., & Morris, S. R. (2007). Clinical implications of the effects of lexical aspect and phonology on children's production of the regular past tense. *Child Language Teaching and Therapy, 23,* 287–306. https://doi.org/10.1177/0265659007080682
Kahn, D. (1976). *Syllable-based generalizations in English phonology* (Doctoral thesis, Massachusetts Institute of Technology). Retrieved from https://hdl.handle.net/1721.1/16397
Kolson, C. (1960). *The vocabulary of kindergarten children* (Unpublished doctoral dissertation), University of Pittsburgh, Pittsburgh, PA.
Leonard, L. B. (2014). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.
Leonard, L. B., Davis, J., & Deevy, P. (2007). Phonotactic probability and past tense use by children with specific language impairment and their typically developing peers. *Clinical Linguistics & Phonetics, 21,* 747–758. https://doi.org/10.1080/02699200701495473
Levelt, C. C., Schiller, N. O., & Levelt, W. J. (2000). The acquisition of syllable types. *Language Acquisition, 8,* 237–264.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38. https://doi.org/10.1017/S0140525X99001776
Lipski, J. M. (1999). The many faces of Spanish /s/-weakening: (Re)alignment and ambisyllabicity. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics* (pp. 198–213). Somerville, MA: Cascadilla Press.
Lleó, C., & Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language, 23,* 31–56.
MacWhinney, B. (2000). *The CHILDES project: Tools for Analyzing Talk* (3rd ed., Vol. II). Mahwah, NJ: Erlbaum.
Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology, and neighborhood structure. *Cognitive Science, 21,* 283–304. https://doi.org/10.1207/s15516709cog2103_2
Marchman, V. A., Wulfeck, B., & Weismer, S. E. (1999). Morphological productivity in children with normal language and SLI: A study of the English past tense. *Journal of Speech, Language, and Hearing Research, 42,* 206–219. https://doi.org/10.1044/jslhr.4201.206
Marinis, T., & Chondrogianni, V. (2010). Production of tense marking in successive bilingual children: When do they converge with their monolingual peers? *International Journal of Speech-Language Pathology, 12,* 19–28.
Marshall, C. R., & van der Lely, H. K. J. (2007). The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Advances in Speech Language Pathology, 9,* 191–203.
Martínez-Gil, F. (2000). La estructura prosódica y la especificación vocálica en español: El problema de la sinalefa en ciertas variedades de la lengua coloquial contemporánea [Prosodic structure and vowel specification in Spanish: The problem of synaloepha in certain varieties of contemporary colloquial language]. In J. G. Fernández (Ed.), *Panorama de la fonología española actual* (pp. 511–560). Madrid, Spain: Arco Libros.
McCarthy, J. J., & Prince, A. S. (1995). Faithfulness and reduplicative identity. In J. Beckman, L. W. Dickey, & S. Urbanczyk (Eds.), *University of Massachusetts occasional papers in linguistics 18: Papers in optimal theory* (pp. 249–384). Amherst, MA: Graduate Linguistic Students' Association, University of Massachusetts.
McLeod, S., van Doorn, J., & Reed, V. A. (2001). Normal acquisition of consonant clusters. *American Journal of Speech-Language*

Pathology, 10, 99–110. https://doi.org/10.1044/1058-0360(2001/011)
Miller, J. F., & Iglesias, A. (2004). Systematic Analysis of Language Transcripts (Version 8), English & Spanish [Computer software]. Madison, WI: Language Analysis Lab, University of Wisconsin.
Moe, A. J., Hopkins, C. J., & Rush, R. T. (1982). *The vocabulary of first-grade children*. Springfield, IL: Charles C Thomas.
Nathan, G. S. (2008). *Phonology: A cognitive grammar introduction*. Amsterdam, the Netherlands: John Benjamins.
Oetting, J. B., & Horohov, J. E. (1997). Past-tense marking by children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 40,* 62–74. https://doi.org/10.1044/jslhr.4001.62
Oetting, J. B., Riviere, A., Lee, R., Richardson, J., Hegarty, M., & McDonald, J. (2012, June). *Structure specific differences in nonmainstream dialects of English across tasks*. Paper presented at the Symposium on Research in Child Language Disorders, Madison, WI.
Owen Van Horne, A. J., & Green Fager, M. (2015). Quantifying the relative contributions of lexical and phonological factors to regular past tense accuracy. *International Journal of Speech-Language Pathology, 17,* 605–616. https://doi.org/10.3109/17549507.2015.1034174
Padilla, A. M. (1978). The acquisition of fourteen English grammatical morphemes in the speech of bilingual children. *Bilingual Review/La Revista Bilingüe, 5,* 163–168.
Paradis, J. (2005). Grammatical morphology in children learning English as a second language. *Language, Speech, and Hearing Services in Schools, 36,* 172–187. https://doi.org/10.1044/0161-1461(2005/019)
Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech, Language, and Hearing Research, 43,* 834–847. https://doi.org/10.1044/jslhr.4304.834
Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders, 43,* 474–497. https://doi.org/10.1016/j.jcomdis.2010.01.002
Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics, 29,* 689–722. https://doi.org/10.1017/S0142716408080296
Pearson, B. Z., Fernandez, S. C., Lewedeg, V., & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics, 18,* 41–58. https://doi.org/10.1017/S0142716400009863
Pitt, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics, 60,* 941–951. https://doi.org/10.3758/bf03211930
Polite, E. J. (2011). The contribution of part-word phonological factors to the production of regular noun plural *-s* by children with and without specific language impairment. *First Language, 31,* 425–441. https://doi.org/10.1177/0142723711406431
Pregibon, D. (1980). Goodness of link tests for generalized linear models. *Journal of the Royal Statistical Society, Series C: Applied Statistics, 29,* 15–23. https://doi.org/10.2307/2346405
Pruitt, S., & Oetting, J. (2009). Past tense marking by African American English–speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 52,* 2–15.
Rice, M. L., Tomblin, J. B., Hoffman, L., Richman, W. A., & Marquis, J. (2004). Grammatical tense deficits in children with SLI and nonspecific language impairment. *Journal of Speech, Language, and Hearing Research, 47,* 816–834. https://doi.org/10.1044/1092-4388(2004/061)
Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech and Hearing Research, 39,* 1239–1257. https://doi.org/10.1044/jshr.3906.1239
Riches, N. (2015). Past tense -ed omissions by children with specific language impairment: The role of sonority and phonotactics. *Clinical Linguistics & Phonetics, 29,* 482–497. https://doi.org/10.3109/02699206.2015.1027832
Rogers, W. (1993). Regression standard errors in clustered samples. *Stata Technical Bulletin, 13,* 19–23.
Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale–Revised (Leiter-R) manual*. Wood Dale, IL: Stoelting.
Rose, Y., & Hedlund, G. (2016). Phon (Version 2.1) [Computer software]. Retrieved from https://www.phon.ca/phontrac
Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). *Akaike information criterion statistics*. Dordrecht, the Netherlands: D. Reidel.
Selkirk, E. O. (1981). English compounding and the theory of word structure. In M. Moortgat, H. van der Hulst, & T. Hoekstra (Eds.), *The scope of lexical rules* (pp. 229–277). Dordrecht, the Netherlands: Foris.
Selkirk, E. O. (1982). The syllable. In H. van der Hulst & N. Smith (Eds.), *The structure of phonological representations* (Vol. 2, pp. 337–383). Dordrecht, the Netherlands: Foris.
Shriberg, L. D., & Austin, D. (1998). Comorbidity of speech-language disorders: Implications for a phenotype marker for speech delay. In R. Paul (Ed.), *The speech-language connection* (pp. 73–117). Baltimore, MD: Brookes.
Song, J. Y., Sundara, M., & Demuth, K. (2009). Phonological constraints on children's production of English third person singular *-s*. *Journal of Speech, Language, and Hearing Research, 52,* 623–642.
Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires: A parent-completed, child-monitoring system* (3rd ed.). Baltimore, MD: Brookes.
Stemberger, J. P. (2007). Children's overtensing errors: Phonological and lexical effects on syntax. *Journal of Memory and Language, 57,* 49–64.
Storkel, H. L., & Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods, 42,* 497–506. https://doi.org/10.3758/BRM.42.2.497
Sundara, M., Demuth, K., & Kuhl, P. K. (2011). Sentence-position effects on children's perception and production of English third person singular -s. *Journal of Speech, Language, and Hearing Research, 54,* 55–71. https://doi.org/10.1044/1092-4388(2010/10-0056)
Szigetvári, P. (2001). Dismantling syllable structure. *Acta Linguistica Hungarica, 48,* 155–181.
Treiman, R., & Zukowski, A. (1990). Toward an understanding of English syllabification. *Journal of Memory and Language, 29,* 66–85. https://doi.org/10.1016/0749-596x(90)90010-w
Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does *not* do). *Journal of Memory and Language, 72,* 37–48.
Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 161–194). Cambridge, United Kingdom: Cambridge University Press.

Chapter 4, in full, is a reprint of material as it appears in Combiths, P., Barlow, J. A., Potapova, I., & Pruitt-Lord, S. (2017). Influences of phonological context on tense marking in Spanish–English dual language learners. *Journal of Speech, Language, and Hearing Research, 60*(8), 2199–2216. The dissertation author was the primary investigator and author of this paper.

CHAPTER 5:

Treatment Targets for Co-Occurring Speech-Language Impairment: A Case Study

Research Article

Treatment Targets for Co-Occurring Speech-Language Impairment: A Case Study

Philip N. Combiths,[a] Jessica A. Barlow,[b] Jennifer Taps Richard,[c] and Sonja L. Pruitt-Lord[d]

Purpose: The intersection of speech and language impairments is severely understudied. Despite repeatedly documented overlap and co-occurrence, treatment research for children with combined phonological and morphosyntactic deficits is limited. Especially, little is known about optimal treatment targets for combined phonological–morphosyntactic intervention. We offer a clinically focused discussion of the existing literature pertaining to interventions for children with combined deficits and present a case study exploring the utility of a complex treatment target in word-final position for co-occurring speech and language impairment.
Method: Within a school setting, a kindergarten child (aged 5;2 [years;months]) with co-occurring phonological disorder and developmental language disorder received treatment targeting a complex consonant cluster in word-final position inflected with 3rd-person singular morphology.
Results: For this child, training a complex consonant cluster in word-final position resulted in generalized learning to untreated consonants and clusters across word positions. However, morphological generalization was not demonstrated consistently across measures.
Conclusions: These preliminary findings suggest that training complex phonology in word-final position can result in generalized learning to untreated phonological targets. However, limited improvement in morphology and word-final phonology highlights the need for careful monitoring of cross-domain treatment outcomes and additional research to identify the characteristics of treatment approaches, techniques, and targets that induce cross-domain generalization learning in children with co-occurring speech-language impairment.

Two of the most prevalent communication disorders in preschool and early school-age children are phonological disorder (PD), a type of speech sound disorder (Law, Boyle, Harris, Harkness, & Nye, 2000), and developmental language disorder (DLD), a type of language impairment also referred to as *specific language impairment* (Bishop, Snowling, Thompson, Greenhalgh, & CATALISE-2 Consortium, 2017). Both PD and DLD are of unknown etiology, occurring independently of a primary motivating condition, and both functionally impair the development, manipulation, and production of the linguistic units of language. The primary distinction between PD and DLD lies in the types of units affected. In PD, phonological units (e.g., speech sounds and sequences) are impacted, whereas in DLD, the impact is seen in larger units (e.g., morphemes, words, and phrases). Despite affecting different language components, PD and DLD both prevent the development of a child's linguistic skills in a timely manner, which can impact their communication and literacy skills and later academic, socioemotional, and occupational outcomes (e.g., Beitchman, Wilson, Brownlie, Walters, Inglis, et al., 1996; Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996).

Because PD has typically been associated with the domain of speech and DLD has been associated with the domain of language, the assessment and treatment of each have most frequently been investigated independently. However, there is a body of evidence that links PD and

[a]San Diego State University/University of California, San Diego
[b]School of Speech, Language, and Hearing Sciences, San Diego State University, California
[c]San Diego Unified School District, California
[d]School of Speech, Language, and Hearing Sciences, San Diego State University, California
Correspondence to Philip Combiths: pcombiths@sdsu.edu

Editor: Brenda Beverly
Received June 28, 2018
Revision received November 28, 2018
Accepted January 22, 2019
https://doi.org/10.1044/2019_PERS-SIG1-2018-0013

Disclosures
Financial: Philip N. Combiths has no relevant financial interests to disclose. Jessica A. Barlow has no relevant financial interests to disclose. Jennifer Taps Richard has no relevant financial interests to disclose. Sonja L. Pruitt-Lord has no relevant financial interests to disclose.
Nonfinancial: Philip N. Combiths has no relevant nonfinancial interests to disclose. Jessica A. Barlow has no relevant nonfinancial interests to disclose. Jennifer Taps Richard has no relevant nonfinancial interests to disclose. Sonja L. Pruitt-Lord has no relevant nonfinancial interests to disclose.

DLD and demonstrates the considerable comorbidity of these disorders. Investigators in the areas of PD (Paul & Shriberg, 1982; Rvachew, Gaines, Cloutier, & Blanchet, 2005) and DLD (Bishop & Edmundson, 1987; Botting & Conti-Ramsden, 2003; Haskill & Tyler, 2007; van Daal & van Balkom, 2004) have repeatedly documented co-occurring deficits, such that difficulty with language (i.e., morphosyntax) is frequently reported in children with PD, and children with DLD often have co-occurring speech-sound production deficits.

Prevalence estimates for co-occurring PD and DLD (hereafter, *PD-LD*) vary greatly. For instance, Shriberg and Kwiatkowski (1994) found between 10% and 77% of 3- to 6-year-old children with speech production deficits demonstrate expressive language concerns (e.g., low utterance length, grammatical morpheme errors, limited vocabulary). In two epidemiological studies reviewed by Pennington and Bishop (2009), children with DLD were 3.3 (Shriberg, Tomblin, & McSweeny, 1999) or 6.1 (Beitchman, Nair, Clegg, Ferguson, & Patel, 1986) times more likely to also have PD. Although these estimates vary, the substantial co-occurrence of phonological and morphosyntactic impairments is well documented, which muddies the distinction between traditional speech and language disorders and demands that this co-occurrence be addressed in the clinical literature.

Converging evidence also suggests that PD and DLD may, at least in some cases, have similar or overlapping etiologies (e.g., Pennington & Bishop, 2009). PD and DLD are associated with multiple cognitive–linguistic deficits, including some that regularly occur in children with either disorder, such as impaired phonological working memory (Archibald & Joanisse, 2009; Raitano, Pennington, Tunick, Boada, & Shriberg, 2004). Heritability has been demonstrated for PD and DLD (Pennington & Bishop, 2009) and is especially linked to instances of PD-LD (Lewis, Freebairn, & Taylor, 2000b). Some have suggested that, in children with PD-LD, low intelligibility characteristic of PD is often identified before characteristics of DLD (e.g., Panagos, 1974), whose presentation is known to change across the life span (e.g., Conti-Ramsden, St Clair, Pickles, & Durkin, 2012). The literature also suggests that children with PD-LD are more severely impaired and are more likely to experience lasting cross-domain deficits than those with PD or DLD alone (Bishop & Edmundson, 1987; Haskill & Tyler, 2007; Ingram, 1976; Lewis, Freebairn, & Taylor, 2000a; Macrae & Tyler, 2014). Furthermore, children with co-occurring PD-LD also demonstrate poorer language and literacy outcomes in adolescence than those with non-comorbid PD or DLD (Lewis et al., 2015).

Overlapping Versus Co-Occurring Disorders

The issue at hand is a group of children with deficits distributed across the nonmutually exclusive domains of phonology and morphosyntax. Notably, not all children with overlapping morphophonological error patterns fit the profile of a child with co-occurring PD-LD. For instance, a child might present with DLD, whose omission of word-final grammatical morphemes, such as third-person singular (3S) *–s* or past tense *–ed*, mimics a pattern of final consonant deletion or cluster reduction (e.g., *goes* /goʊz/ as [goʊ], *stopped* /stɑpt/ as [stɑp]), which is a relatively common phonological process in children with PD (Ingram, 1989). This scenario could presumably be identified as a case of DLD—and not PD—by examining word-final production of [–s], [–z], [–t], and [–d] in nonmorphemic word-final contexts. A notably lower marking rate for 3S [–z] in a word such as *goes* /goʊz/ than for nonmorphemic [–z] in a word such as *hose* /hoʊz/ would suggest difficulty with production of the grammatical marker, unrelated to production of word-final [–z].

Alternatively, a child might present with PD, whose limited or absent production of phonemes in coda position, certain consonants, or consonant clusters results in limited or absent marking of word-final grammatical morphemes. For instance, a child who does not produce [s] in words of any type would also omit 3S [–s] in *she walks*; however, in this instance, the underlying deficit would lie in phonology and not in morphosyntax. In this case, a thorough examination of the child's phonological system, especially productions of [–s], [–z], [–t], and [–d] in word-final position, would permit identification of phonological constraints impacting morpheme production, and this could be compared to morphosyntactic knowledge in less phonologically constrained contexts, such as vowel-final verbs (e.g., *goes* /goʊz/) or other non–word-final tense and agreement morphemes, such as copula or auxiliary *BE*.

A third profile would be that of a child with PD-LD—co-occurring phonological and morphosyntactic deficits above and beyond surface-level overlap in cases where these two domains intersect. A child with co-occurring PD-LD would demonstrate morphological deficits beyond contexts where phonology would impact morphology and vice versa. These children would demonstrate deficits in both morphosyntax and phonology in contexts other than their overlap in word-final position, such as errored case marking and low word-initial consonant accuracy. Given that PD, DLD, and PD-LD can resemble each other when viewed through certain lenses, thorough cross-domain assessment should be conducted to reduce the risk of an incomplete or inappropriate diagnosis.

Cross-Domain Treatment: Approaches, Techniques, and Targets

Several studies have documented approaches that resulted in quantifiable improvement in both phonology and morphosyntax following intervention for PD-LD. Cross-domain improvements were documented in one study that sequenced or alternated treatment domains (Tyler, Lewis, Haskill, & Tolbert, 2003) and in three studies that combined phonological and morphological treatment in the same session or activity (Bellon-Harn, Hoffman, & Harn, 2004; Tyler et al., 2003; Tyler & Sandoval, 1994).

Interestingly, several studies also reported cross-domain improvement following morphosyntactic treatment only (Hoffman, Norris, & Monjure, 1990, 1996; Matheny & Panagos, 1978; Tyler et al., 2003) or phonological treatment only (Matheny & Panagos, 1978). Additional studies using combined treatment approaches for this population only found well-documented improvement in one of the two linguistic domains (i.e., phonology: Hodson, Nonomura, & Zappia, 1989; morphosyntax: Fey et al., 1994; Tyler, Lewis, Haskill, & Tolbert, 2002). Consequently, cross-domain improvement may be achievable for children with PD-LD, although methodological variations across studies make it difficult to determine which domain(s) should be targeted for optimal results.

Perhaps, less clear to a clinician faced with providing intervention for a child with PD-LD are which techniques to use or which linguistic structures to target in intervention. In the aforementioned studies, a variety of intervention techniques were successfully employed, including focused stimulation or drill, minimal pair contrasting, elicited production, cloze techniques (Tyler et al., 2002, 2003), phonological and auditory awareness activities (Gillon, 2000; Tyler et al., 2003), naturalistic play, and storytelling (Hoffman et al., 1990, 1996). Although no single approach has emerged as most effective, fortunately, the implication is that a variety of intervention activities can achieve cross-domain improvement.

There is especially little evidence as to the characteristics of the linguistic structures that should be targeted in treatment for children with PD-LD. In several of the studies reviewed above, phonological or morphosyntactic targets were described only broadly as complex discourse, narrative structures, or phonological processes (Bellon-Harn et al., 2004; Hoffman et al., 1996). In the work by Tyler and colleagues (Tyler et al., 2002, 2003; Tyler & Sandoval, 1994), phonological processes and morphosyntactic errors were directly targeted by training specified linguistic structures. For example, in Tyler et al. (2002), treatment for one child targeted /fl–/ and /kl–/ to address initial cluster reduction and targeted past tense *–ed*, irregular past tense, possessive *'s*, and auxiliary *BE* to address grammatical morphology errors. In each of these studies, treatment targets were scheduled using a cycles approach (Hodson & Paden, 1991), which necessarily involves the use of multiple targets. However, this limits our ability to retrospectively identify the influence of a given linguistic target (e.g., /fl–/ or past tense *–ed*), as it cannot be separated from the influence of other simultaneous or cycled targets.

Treatment Target Complexity and Generalization

Because treatment research targeted for children with PD-LD is still relatively sparse and the studies that do exist have (a) broadly targeted one or more domains, (b) not specified treatment targets, or (c) employed multiple simultaneous or cycled targets, attention has yet to shift to the impact of characteristics of the stimuli or treatment targets that are trained in combined phonological–morphosyntactic interventions. This is an important consideration for effective intervention because of the impact that a given treatment target or stimulus may have on generalization learning (Cummings & Barlow, 2011; Gierut & Morrisette, 2010; Plante et al., 2014; Van Horne, Fey, & Curran, 2017).

Introducing linguistic structures into an impaired system by training linguistically complex exemplars is an effective approach for treatment of disorders in multiple linguistic domains (e.g., syntax: Thompson, Shapiro, Kiran, & Sobecks, 2003; morphology: Van Horne, et al., 2017; phonology: Gierut, 2008). Per this complexity-based approach, new language structures are learned by exposure to input that demonstrates a greater extent of higher level variations. Presumably, relatively complex exemplars in the input require the child's existing language knowledge to expand in accommodation of the new, complex input (Gierut, 2007). In practice, the primary advantage of a complexity-based approach is that it is expected to result in simultaneous generalization learning of simpler, untreated components of the language (e.g., Storkel, 2018a). Within the phonological domain, teaching complex sound targets results in greater generalized learning across sound manner classes (e.g., stops, fricatives, nasals) than the traditional approach of teaching simpler targets first (e.g., Tambyraja & Dunkle, 2014). Furthermore, complex consonant cluster targets, such as /spl–/ in *splash,* appear to result in the greatest amount of generalization learning, improving accuracy of other consonant clusters and singleton consonants and, consequently, resulting in higher consonant accuracy and overall intelligibility (Elbert, Dinnsen, & Powell, 1984; Gierut, 1999; Gierut & Champion, 2001).

Within the domain of morphosyntax, treatment using complex stimuli has been shown to result in improved generalization learning in adults with aphasia at the sentence level (e.g., Thompson, Ballard, & Shapiro, 1998) and children with DLD at the morphological level (Van Horne et al., 2017). In their recent study, Van Horne et al. (2017) targeted past tense morphology in children with DLD using a set of verb stimuli classified as difficult or easy based on their phonological, semantic, and frequency characteristics. Children trained first using the more difficult and ostensibly more complex stimuli made greater improvement, including generalized learning to untreated verbs, than the children trained first with the easier, or simpler, stimuli. Strategic manipulation of the complexity of treatment targets and stimuli has thus been associated with improved generalization learning, not only for children with PD but also for individuals with language impairments, including DLD.

Complex Treatment Targets for PD-LD

The apparent effectiveness of complex consonant clusters in word-initial position for treatment of PD leads us to consider how complex treatment targets can be applied to children with co-occurring PD-LD. In English, there is overlap between complex phonology and morphology in word-final verb tense markers (e.g., 3S *–s* and past tense *–ed*). For example, suffixation of [–s] or [–z] in a

3S-inflected verb not only adds morphological content to the verb but also increases its phonological complexity by adding a coda consonant (e.g., *fly+s* /flaɪz/) or forming a consonant cluster (e.g., *eat+s* /its/, *help+s* /hɛlps/).

To our knowledge, there has been no isolated investigation of treatment outcomes for children with PD-LD as a result of training a complex morphophonemic consonant cluster (e.g., a tense-inflected word-final cluster). There is, however, some precedent for targeting morphophonemic word-final clusters in this population. In their treatment study, Tyler and Sandoval (1994) targeted the phonological processes of final consonant deletion or cluster reduction in six preschool-age children with PD-LD, with the expressed intention of improving the production of grammatical morphemes that occur in word-final consonants and clusters. The children in their study received one of three interventions: (a) a phonological intervention targeting phonological processes using elicited imitation and minimal pair contrasts, (b) a narrative-based intervention targeting various grammatical morphemes and expanded utterances using focused stimulation and recasting, or (c) a combination of both interventions in the same session.

Because the interventions in Tyler and Sandoval (1994) involved a cycles approach to treatment targets, the independent effectiveness of targeting complex word-final phonology and morphology together is impossible to determine in this study. Notwithstanding, two participants, S1 and S6, were treated by expressly targeting word-final cluster reduction in morphologically complex contexts in one of their cycles. Participant S1 received the direct phonology treatment and was trained to produce /–ps/ and /–ts/ to target final cluster reduction among his other cycled targets. Following 12 weeks of treatment, S1's frequency of occurrence of final cluster reduction was reduced from 100% at baseline to 21%. His production accuracy of 5 two-element final clusters (including the two treated clusters) increased from 0% at baseline to 58%.

Participant S6, who received the combined phonological and narrative-based intervention, was similarly trained to produce /–ts/ to target final cluster reduction among his other cycled targets. Following 12 weeks of treatment, S6's frequency of occurrence of final cluster reduction was reduced from 100% at baseline to 0%. His production accuracy of 5 two-element final clusters (including the one treated cluster) increased from 0% at baseline to 100%. Notably, S1 and S6 outperformed the other four children who were not trained on word-final consonant clusters in terms of overall consonant accuracy improvement and elimination of phonological processes, although they also had greater receptive language skills, percentage of consonants correct, and mean length of utterance (MLU) at baseline than other participants.

Although the impact of treatment targeting final cluster reduction and, indirectly, complex final clusters in Tyler and Sandoval's (1994) study cannot be isolated from the impact of other cycled targets, the children who were trained to produce two-element morphophonemic final clusters as part of their intervention outperformed many of the other participants following treatment, which suggests that targeting morphologically inflected final clusters might be beneficial for children with PD-LD and warrants further investigation.

A Case Study

Following these suggestive results, we will directly explore the use of a single morphophonologically complex treatment target by describing the execution and outcome of a school-based case study conducted with a young child with PD-LD. The treatment target and treatment protocol will be described, and treatment outcomes will be presented and discussed as they relate to morphophonological learning and the variable profiles of children presenting with overlapping characteristics of PD and LD or co-occurring PD-LD.

The study participant was a child in kindergarten, herein referred to as "Max," who demonstrated a profile of co-occurring PD-LD. At the beginning of the school year, Max presented with spoken language deficits in the areas of phonology and morphosyntax, including—but crucially not limited to—overlapping errors in phonology and morphology in word-final position (i.e., final consonant deletion/cluster reduction and word-final grammatical morpheme omission). Although Max presented with an overlapping deficit pattern that could be associated with either PD or LD independently, his phonological and morphosyntactic deficits were not exclusive to this overlap, as described below. Thus, Max had co-occurring PD-LD, rather than limited overlapping features of LD ultimately attributable to PD.

A speech-language pathologist (SLP) faced with a child, such as Max, with PD-LD must choose treatment approach (es) and target(s) to address both of the child's impacted language domains. Should an SLP decide to approach treatment of phonology and morphosyntax simultaneously, there is little guidance as to effective targets to train during treatment. Max's treatment presented here is a direct response to the scarcity of treatment target efficacy evidence available to guide target selection for SLPs treating children with PD-LD. Furthermore, Max's treatment was conducted as it would occur in typical school-based services; intervention was provided in a small group with another child of similar age, occurred on-site at his elementary school, and was scheduled according to his Individualized Education Program (IEP) requirements and academic schedule.

Specifically, Max received 14 weeks of direct phonological treatment targeting one word-final consonant cluster in real and nonce stimulus verbs inflected with a 3S morpheme (e.g., *he helps*). The goal of treatment target selection was to identify a linguistically complex structure, /–lps/, with phonological and morphological content whose accuracy was low at baseline. With this complex target, we expected to observe generalization learning beyond the treated consonant cluster in Max's language production via his phonological system, his morphological system, or both. The impact of this approach was determined by posttreatment generalization learning of phonological and morphological structures as measured by singleton consonant, cluster, and grammatical morpheme accuracy outside treated contexts.

Method

Participant

Max was a monolingual, English-speaking child (male, aged 5;2 [years;months]) attending a public kindergarten. In the summer prior to this study, he received an IEP identifying speech and language concerns and participated in a preschool speech improvement class (a general education service; Montgomery, Dunaway, & Taps, 2005) targeting structured conversation, including words with initial /fl–/ and /skw–/ sounds. However, upon entry to kindergarten in the fall, he received no speech-language services other than the intervention described here.

Max's speech-language skills were assessed using a battery of elicitation probes (described below) and a language sample collected at baseline. Max's elicited productions were characterized by low consonant accuracy and notable word-final morpheme omissions, including 0% marking of regular monosegmental 3S, past tense, and plural morphemes. His pretreatment singleton and cluster accuracy as well as word-final morpheme marking rates from baseline probes are shown in Table 1. His pretreatment phonetic inventory of singleton consonants and clusters in word-initial and word-final positions from these same probes is given in Table 2. Specifically, Max's percentage of consonants correct-revised (PCC-R) was 43.3%. Speech-sound production accuracy was especially low for word-initial clusters (25.7%) and singletons and clusters in final position (31.2% and 7.6%, respectively), and he did not mark (i.e., produce) any instances of word-final [–s, –z, –t, –d] morphemes in elicitation probes. His inventory of singletons and clusters (see Table 2) was also limited, especially in word-final position—with notable absence of final [–s] and [–z], which are used for regular 3S and plural marking.

Phonological and morphosyntactic characteristics of Max's elicited productions were similar to those attested in his conversational connected speech. Per a 152-utterance language sample, Max presented with highly unintelligible speech and low consonant accuracy in addition to notable morphosyntactic deficits, including 0% regular 3S and past tense marking, and frequent omission of plural marking, copula and auxiliary *BE*, and functor words (e.g., *to*) and incorrect case marking (e.g., *him* for *he*). His MLU in morphemes was 3.18.

Table 1. Baseline percent correct singletons and clusters by position, percentage of consonants correct–revised (PCC-R), and monosegmental regular grammatical morpheme marking rate.

	Phonology, % correct			Morphology, % marked	
Measure	*Word-initial*	*Word-final*	*All positions*	Morpheme	%
Singletons	62.8	31.2	47.4	3S	0.0
Clusters	25.7	7.6	15.3	Past tense	0.0
PCC-R	—	—	43.3	Plural	0.0

Note. Em dash indicates values were not calculated.

Max's phonological characteristics, including highly unintelligible speech and low consonant accuracy, are indicative of moderate-to-severe PD, and his omission of grammatical morphemes, including those that are not restricted by complex phonology (e.g., auxiliary and copula *BE*, articles, and prepositions), is consistent with a profile of DLD. Per his initial IEP report, based primarily on performance-based assessment and parent report, Max did not have any other co-occurring motor, social–emotional, behavioral, or health concerns. Max also passed a binaural hearing screening at the beginning of the school year. Together, these phonological and morphosyntactic deficits in the absence of other concomitant impairment indicate a co-occurring PD-LD profile.

All baseline/generalization probes and intervention sessions, as described herein, were administered by a certified SLP or graduate student clinician under the former's supervision. All data collection and treatment were conducted on-site at Max's school.

Generalization Probes

A set of morphophonological generalization probes was administered prior to treatment to establish baseline performance, as described above, and at time points during and after treatment. A complete probe administration occurred before and after each treatment phase and after each break during which no treatment occurred, as shown in Table 3. A subset of the generalization probe administered at baseline was administered twice to establish baseline stability from September to mid-October.

The probe battery consisted of the singletons and initial clusters in single-word elicitation probes of the In-Depth Phonological Assessment (Taps Richard, 2012) and a word-final probe sampling final clusters and regular word-final morphology (plural, 3S, and past tense) in a variety of word-final phonological contexts (i.e., preceded by different types of consonants, clusters, and vowels). Word lists for the generalization probe are available in the Appendix. The complete battery elicited 312 words and sampled every singleton consonant at least three times in initial, medial, and final positions; 28 initial clusters and 30 final clusters were sampled at least twice. Each grammatical morpheme was sampled, at minimum, 13 times in its monosegmental form (i.e., /–s/, /–z/, /–t/, /–d/) across a variety of phonological contexts.

Elicitation probes were administered from a notebook or folder with images (four per page) or a tablet screen (one image at a time). Max's productions were elicited without a model whenever possible; however, when he did not know the target or produced a nontarget word, the target was elicited via delayed imitation of the examiner's model: "Is it ___ [target word] or ___ [unrelated word]?" or "This is a ___ [target word]. Tell me what this is." Word-final

Table 2. Baseline phonetic inventory of singletons and clusters, by word position.

Word position	Singletons	Clusters
Word-initial	[b d h j k l m n p ɹ s t ʧ w]	[bl bɹ bw gɹ kl kɹ pl pɹ pw tl tɹ tw]
Word-final	[b d k l m n ŋ p ɹ t v ʔ]	[lp mp nt ɹm ɹn ɹt ɹʔ]

tense morpheme production was elicited using cues adapted from the 3S and past tense probes of the Rice/Wexler Test of Early Grammatical Impairment (Rice & Wexler, 2001), with a single practice item preceding each block of 3S or past tense elicitations. These elicitations were in the following form: "Look at this woman. Tell me what she does: she ___ [3S target]" or "She's jumping. Now she's done. Tell me what she did: she ___ [past tense target]."

Treatment Target and Stimuli

Max's treatment target was selected by adapting procedures for selecting complex word-initial phonological treatment targets (Morrisette, Farris, & Gierut, 2006). These procedures take into account child-specific and linguistic factors that affect the complexity and generalizability of potential treatment targets, including the child's knowledge of each component segment (i.e., accurate or inaccurate use and stimulability) and the linguistic complexity of the target. For additional information related to using complex phonological targets in treatment, readers are referred to Storkel (2018a). In adapting these procedures for a morphophonemic word-final target, a three-element cluster was selected because of its phonological complexity and the absence of any three-element clusters in Max's pretreatment inventory. The target exemplar, /–lps/, was selected because of the complexity of these sounds in combination and Max's poor pretreatment knowledge of the component sounds in word-final position. Finally, the 3S morpheme was selected rather than the plural morpheme because 3S is generally acquired later (Brown, 1973) and is typically more severely impacted in children with DLD (Leonard, 2014).

Treatment stimuli were a set of seven verb phrases (four real and three nonce verbs). Both types of verbs were selected because nonwords have been shown to induce generalization learning more readily than real words in children with PD (Cummings & Barlow, 2011; Gierut & Morrisette, 2010; Storkel, 2018b). However, less is known about the impact of nonwords—particularly nonce verbs—in treatment for children with PD-LD; thus, both were included in Max's set of stimuli. For the first phase of treatment, stimulus verbs were preceded by the third-person pronoun *he* to create short, grammatically correct phrases (e.g., *he helps*, *he malps*) appropriate for imitation and drill-based activities, although the subject agent was varied during the second phase of treatment. All consonants in stimulus phrases, except for the cluster target, were attested in the child's pretreatment phonetic inventory to encourage focused learning of the targeted novel cluster. The stimulus phrases used in this study are given in the Appendix.

Table 3. Schedule of baseline/generalization probes, periods of treatment, and breaks during which no treatment occurred.

Month		Probe/treatment
September		Baseline probes
	Treatment Phase I	
December		Post–Phase I probes
	Break 1	
January		Post–Break 1 probes
	Treatment Phase II	
March		Post–Phase II probes
	Break 2	
April		Post–Break 2 probes

Treatment Protocol

Max received 14 weeks of treatment provided in 30-min small-group sessions twice per week, as is typical for school-based services. The small group consisted of the participant and one other child receiving treatment with a nonmorphemic phonological target. The treatment followed a drill-play format (Shriberg & Kwiatkowski, 1982) with a production phase (Phase I) followed by an extension phase (Phase II). During the production phase, Max was introduced to four of the treatment stimuli and began producing target words following the clinician's verbal model with 1:1 clinician feedback. Explicit articulatory instruction was introduced and used as needed, including verbal, tactile, and kinesthetic cues, to elicit correct target forms (Bauman-Waengler, 2008; Secord, 2007). Max successfully imitated the target cluster across multiple attempts by the end of the first week of treatment; thus, for the remainder of the production phase, Max produced target verb phrases spontaneously or through elicitation but without an immediate verbal model. As the production phase progressed, corrective feedback was intermittently delayed or withheld to provide opportunities for self-monitoring and self-correction (Ertmer & Ertmer, 1998; Shriberg & Kwiatkowski, 1990). The role of 3S *–s* was not explicitly instructed during this phase. Treatment materials included toys and picture cards depicting target verbs. Activities used to generate target productions included matching, sorting, and structured play. Max produced the treatment target, on average, 81 times per session (range: 24–150) across both phases of treatment. The end of Phase I (Week 8) coincided with the beginning of an academic break during which 5 weeks of no treatment occurred. Any change in speech-sound or word-final morpheme production during this period of no treatment was monitored by administering the generalization probe before and after the break.

During the extension phase (Phase II), the same phonological sequence was targeted; however, three additional treatment verbs (one real and two nonce verbs) were introduced, and production opportunities were extended to the sentence, conversation, and narrative levels through role play and collaborative storytelling. This phase continued to target accurate production of the target cluster; however, the role of 3S *–s* was explained explicitly at the start of the extension phase, and Max was encouraged to combine treatment verbs with other 3S subjects (e.g., *Laura gulps*; *the dog telps*). The conclusion of phonological treatment (Week 14) also coincided with the beginning of an academic break during which 5 weeks without treatment occurred. As before, change occurring during the 5-week period after treatment was monitored by performance on the generalization probe.

During both phases of intervention, progress was monitored using a 1–7 visual analog scale (Munson, Schellinger, & Urberg Carlson, 2012) that evaluated production of the trained three-element word-final cluster in treatment words. Each treatment target word was elicited once at the end of treatment sessions with no clinician model or feedback. Max's production of the target cluster was evaluated from 1 (*only a single segment, not in target cluster*) to 7 (*three correct target segments*). A sample rating scale can be found in Figure 1.

Outcome Measures

Phonetic transcriptions of baseline and generalization probes were the primary source of data for the outcome measures in this study. All probes were transcribed online by the administering clinician or supervised graduate student using narrow phonetic transcription and audio-recorded for later transcription by a research assistant trained in phonetic transcription of child speech to determine transcription reliability. Audio recordings were consulted for any words that were unclear to the online transcriber to amend missing or incomplete transcriptions. Transcription reliability for consonants, including word-final morpheme segments, was calculated for 30% (662 words) of the transcribed generalization probes at 85% (Shriberg & Lof, 1991).

Treatment outcome was defined as generalization learning in the phonological and morphological domains in nontreated contexts. Because phonological generalization from a complex target is expected to occur across contexts and sound classes, Max's outcome measure in the phonological domain was defined as production accuracy for all singletons and clusters in nontreated words that were produced with 0% accuracy prior to treatment. An additional clinical measure, PCC-R, and a descriptive measure, phonetic inventory in each word position, were also collected. Max's outcome measure in the morphological domain was defined as regular monosegmental 3S, past tense, and plural marking rate in nontreated words in the generalization probes. Probe transcriptions were analyzed for consonant accuracy, morpheme marking, and PCC-R using Phon (Rose & Hedlund, 2017), a free and open-source software for phonological analysis. Phonetic inventories were collected with the AutoPATT plugin (Combiths, Amberg, & Barlow, 2016) for Phon. Additional language sample measures (MLU in morphemes and grammatical morpheme marking rates) were derived from pre- and posttreatment language samples orthographically transcribed and coded for morpheme use by SLP graduate students using the Systematic Analysis of Language Transcripts software and guidelines (Miller & Iglesias, 2012).

Results

Baseline Stability

In this case study, Max's baseline stability served as his own control. A stable baseline prior to treatment provides a point of comparison by which subsequent changes can be attributed to treatment, rather than maturation.

Figure 1. Sample target production rating scale.

Sample Visual Analog Scale* Target Accuracy Data Collection Form

Treatment Word(s)	Date:	Date:	Date:
	1 2 3 4 5 6 7 ⟷	1 2 3 4 5 6 7 ⟷	1 2 3 4 5 6 7 ⟷
	1 2 3 4 5 6 7 ⟷	1 2 3 4 5 6 7 ⟷	1 2 3 4 5 6 7 ⟷

1 = one sound not in cluster (e.g., [w] for /–lps/)
2 = one sound in cluster (e.g., [p] or [l] for /–lps/
3 = two sounds (not in cluster) (e.g., [wt] for /–lps/)
4 = two sounds (one not in cluster) (e.g., [wp] for /–lps/)
5 = two sounds (in cluster) (e.g., [lp] for /–lps/)
6 = 3 sounds (2 sounds in cluster accurate, other 1 approximation) (e.g., [lps̪] for /–lps/)
7 = three sounds correct (e.g., [lps] for /–lps/)

* Based on the work of Munson, Schellinger, & Urberg Carlson, 2012

Relative stability of his phonological and morphological systems prior to treatment, defined as < 10% change during baseline, was established across a set of probes administered in the month of September (Baseline 1) and another set administered in early–mid October (Baseline 2). Treatment did not commence until after completion of Baseline 2 probes. Singleton and cluster accuracy from an overlapping subset of 54 probe words across staggered pretreatment time points (see Table 4) indicates stability during the 28-day period prior to participation in this study. Regular monosegmental 3S, plural, and past tense morpheme production was also measured and remained stable at 0% across Baseline 1 and Baseline 2 pretreatment probes.

Treatment Progress

Visual analysis of accuracy on the treated target in trained stimulus phrases (see Figure 2) indicated that learning of the targeted morphophonological structure, /–lps/, was occurring during treatment sessions. In-session performance rose to ceiling by Week 14 of treatment, demonstrating Max's progress in trained treatment verbs over the course of the intervention.

Phonology

With treatment targeting a complex word-final morphophonological sequence, we expected generalization learning to occur in Max's phonological system, his morphological system, or both. Composite accuracy scores were calculated for all singletons and clusters that were produced with 0% accuracy prior to treatment (herein referred to as 0%-at-baseline sounds) by word position (initial, final) and type (singleton, cluster) after each treatment phase, as shown in Figure 3. Max's 0%-at-baseline singletons in word-initial position demonstrated the most observable change, improving from 0% to 43% by the end of treatment. Composite accuracy for 0%-at-baseline initial clusters, final singletons, and final clusters improved between 10% and 22% over the course of treatment.

Tables 5 and 6 display accuracy for each 0%-at-baseline singleton or cluster in word-initial and word-final positions, respectively, after Phases I and II of treatment. In word-initial position, six 0%-at-baseline singletons showed improvement following treatment. Of these, [z–], [g–], and [ʃ–] improved the most, with accuracy improving from 0% at baseline to 57%–80% posttreatment. Seven 0%-at-baseline consonant clusters showed improved accuracy following treatment, including [sl–] and [tw–], which rose from 0% to 100% accuracy following treatment.

In word-final position, despite many singletons produced with 0% accuracy prior to treatment, only one of these singletons appears to have been impacted by treatment. Greater improvement was seen in word-final [–l], which improved from 0% to 67% accuracy posttreatment. As with word-final singletons, clusters in word-final position were limited prior to treatment. With treatment, Max demonstrated improved accuracy in eight 0%-at-baseline final clusters, although improvement was generally limited and, in the case of [–lk], [–lps], [–mps], and [–mpt], was not persistent into the second phase of phonological treatment.

Table 4. Percentage of consonants correct-revised (PCC-R) and regular word-final morpheme marking rate stability across baselines.

Measure	Baseline 1	Baseline 2
PCC-R	41.5%	41.7%
3S	0.0%	0.0%
Plural	0.0%	0.0%
Past tense	0.0%	0.0%

Note. 3S = third-person singular.

Figure 2. Production accuracy rating (1–7) of /–lps/ cluster production in target verbs by week.

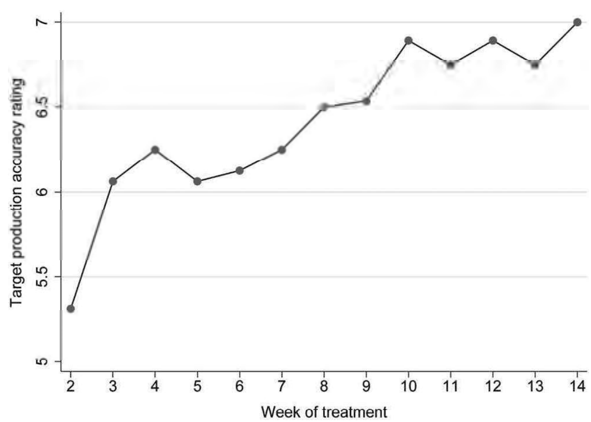


Figure 3. Composite accuracy of 0%-at-baseline singletons and clusters by word position.

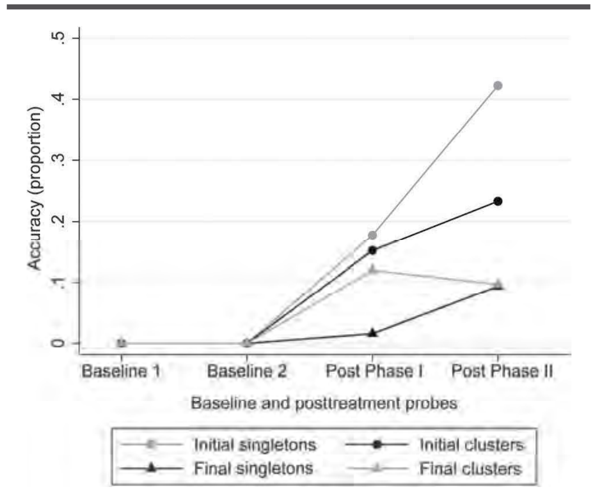


Table 5. Zero-percent-at-baseline word-initial singletons and clusters showing change over the course of treatment.

	Baseline		Post Phase I		Post Phase II	
Sound	Accuracy	CI	Accuracy	CI	Accuracy	CI
ʤ–	0%	—	33%	[1, 66]	11%	[–11, 33]
g–	0%	—	11%	[–11, 33]	67%	[34, 99]
ʃ–	0%	—	29%	[–8, 65]	57%	[17, 97]
v–	0%	—	0%	—	17%	[–16, 49]
z–	0%	—	0%	—	80%	[41, 119]
θ–	0%	—	25%	[–7, 57]	38%	[2, 73]
fl–	0%	—	33%	[–32, 99]	33%	[–32, 99]
fɹ–	0%	—	0%	—	33%	[–32, 99]
kw–	0%	—	33%	[–32, 99]	75%	[26, 124]
sl–	0%	—	100%	—	100%	—
sm–	0%	—	0%	—	33%	[–32, 99]
sw–	0%	—	100%	—	50%	[–48, 148]
tw–	0%	—	33%	[–32, 99]	100%	—

Note. Em dash indicates values were not calculated. CI = 95% confidence interval.

Although Max demonstrated only transient improvement in three-element final clusters and two-element liquid + stop clusters, more persistent improvement occurred within the nasal + stop class of two-element final clusters.

The expansion of Max's phonetic inventory in each word position was also monitored as a descriptive measure of treatment progress. His phonetic inventory by word position following word-final treatment also reflects generalized learning to other untreated consonants and clusters. Inclusion in the phonetic inventory at any time point for a given word position was determined by two or more occurrences at a given probe point. In initial position, Max added four singletons [θ–, ʃ–, f–, and g–] and five clusters [gl–, dɹ–, gɹ–, kw–, and sl–] to his phonetic inventory. In final position, he added four clusters [–lt, –nd, –ɹd, and –ɹp] to his phonetic inventory.

Grammatical Morphology

In addition to targeting the participant's phonological system, the word-final treatment also indirectly targeted word-final grammatical morphology in the form of 3S inflection. Per his performance on the generalization probes (shown in Table 7), Max demonstrated only nominal improvement to regular 3S and past tense and no improvement to plural morphology. However, Max's productions in a 397-utterance posttreatment language sample do suggest growth in his morphological system following treatment (shown in Table 8). Comparing his pre- and posttreatment language samples, Max's MLU in morphemes increased from 3.18 to 3.39, marking rate for regular 3S increased from 0% to 5%, regular past tense increased from 0% to 53%, copula *BE* increased from 8% to 50%, present progressive increased from 10% to 100%, and auxiliary *DO* increased from 57% to 64%. Because baseline stability was only established for performance in the generalization probes—not for language sample measures, Max's apparent improvement in morphological productions from his language sample cannot be attributed to treatment above and beyond maturation.

Table 6. Zero-percent-at-baseline word-final singletons and clusters showing change over the course of treatment.

	Baseline		Post Phase I		Post Phase II	
Sound	Accuracy	CI	Accuracy	CI	Accuracy	CI
–l	0%	—	11%	[–11, 33]	67%	[34, 99]
–lk	0%	—	50%	[–48, 148]	0%	—
–lps	0%	—	33%	[–32, 99]	25%	[–24, 74]
–mps	0%	—	50%	[–7, 107]	0%	—
–mpt	0%	—	50%	[–48, 148]	0%	—
–nd	0%	—	33%	[–8, 75]	33%	[–8, 75]
–ŋk	0%	—	35%	[–24, 74]	75%	[26, 124]
–ɹd	0%	—	40%	[–8, 88]	40%	[–8, 88]
–ɹts	0%	—	33%	[–32, 99]	0%	—

Note. Em dash indicates values were not calculated. CI = 95% confidence interval.

Treatment Impact and Effect Size

Because generalization in the phonological domain was observed following treatment, the impact of intervention on Max's phonological system was further described with two clinically relevant metrics: PCC-R and standard mean difference effect size following Gierut, Morrisette, and Dickinson (2015). These measures were calculated to determine the impact of treatment and quantify change in Max's overall speech-sound production accuracy. Furthermore, PCC-R, as scored for each consonant in each word of the generalization probes, served as a binary outcome measure suitable for logistic regression comparisons of PCC-R before and after each treatment phase and each 5-week break in treatment. Significantly different change during treatment phases in the absence of significant change during breaks without treatment would support the impact of treatment on Max's phonological system, above and beyond identifiable maturation effects.

PCC-R was calculated at each probe administration and compared pre-to-post treatment phases and pre-to-post breaks (shown in Table 9). PCC-R after Phase I of treatment (post Phase I; PCC-R = 55.8%) was significantly higher than at baseline (Baseline; PCC-R = 43.3%, $p < .01$), and PCC-R improved significantly during Phase II of treatment (post Phase II; PCC-R = 60.9%, $p < .01$). Because a young child's sound system can be expected to

Table 7. Regular word-final morpheme marking rates.

Morpheme	Baseline	Post Phase I	Post Phase II
3S	0.0%	6.3%	5.6%
Plural	0.0%	0.0%	0.0%
Past tense	0.0%	0.0%	8.3%

Note. 3S = third-person singular.

Table 8. Morpheme marking rates from pre- and posttreatment language samples.

	Pretreatment		Posttreatment	
Morpheme	% Marked	Opportunities	% Marked	Opportunities
3S	0	3	5	20
Plural	33	3	83	12
Past tense	0	3	53	15
Progressive	10	10	100	25
Aux *DO*	57	7	64	22
Copula *BE*	8	13	50	22
Aux *BE*	33	3	0	15

Note. 3S = third-person singular; Aux = auxiliary.

experience maturational growth over time, pre-to-post treatment differences alone cannot identify a causal impact of word-final phonological treatment on PCC-R improvement. However, PCC-R remained stable at baseline (see Table 4), and PCC-R did not increase during either 5-week break without treatment (see Table 9). This suggests that the improvements in PCC-R after Phases I and II of treatment can be attributed to the effect of word-final phonological treatment and not to maturation.

Treatment effect size was determined using standard mean difference based on change in accuracy of all 0%-at-baseline singletons and clusters following Gierut et al. (2015), calculated as (mean accuracy across all baseline probes) − (mean accuracy across all mid- and posttreatment probes) / (pooled population standard deviation at baseline). Because a population standard deviation cannot be determined for a single child, the population standard deviation from Gierut et al. ($SD = 0.02$) was used as the denominator in the effect size calculation. The effect size of Max's treatment was 7.85—a moderately large effect (large effect range: 6.32–27.83) for treatment of PD according to the benchmarks established from the treatment outcomes of 135 preschool children.

Table 9. Percentage of consonants correct-revised (PCC-R) by time point with logistic regression results.

Time	PCC-R	95% CI	χ^2	p^a
Baseline	43.3%	[40.2, 46.5]	—	—
Post Phase I	**55.8%**	**[52.7, 58.9]**	**29.8**	**< .01***
Post Break 1	53.6%	[50.4, 56.7]	1.0	1.00
Post Phase II	**60.9%**	**[57.8, 63.9]**	**10.5**	**< .01***
Post Break 2	58.3%	[55.2, 61.4]	1.3	1.00

Note. Em dash indicates values were not calculated. χ^2 contrasts compared PCC-R at each posttreatment time point (i.e., post Phase I, post Phase II) to PCC-R prior to that treatment phase (i.e., baseline, post Break 1, post Break 2). χ^2 contrasts also compared pre and post each 5-week period of no treatment (i.e., Break 1 and Break 2). Posttreatment time points are in boldface. CI = confidence interval.

[a]Significant difference in PCC-R between pre- and post-treatment phases.

*Bonferroni-adjusted *p* values.

Discussion

There is a paucity of evidence-based support for SLPs faced with the assessment and treatment of children with overlapping or co-occurring speech and language impairments. Children with PD may present with speech production patterns that resemble morpheme omission errors typical of DLD, such as final consonant deletion or final cluster reduction, and conversely, children with DLD may omit word-final grammatical morphemes, mimicking word-final phonological processes. There is, however, a considerable subgroup of children who present with co-occurring PD-LD beyond the morphophonological overlap that occurs in word-final position. Treatment research for this population has, so far, demonstrated that a variety of combined, sequenced, or alternating intervention approaches and various intervention techniques can lead to improved outcomes in phonology and morphosyntax for these children. However, identification of the treatment target characteristics that impact treatment outcomes for children with PD-LD is only just emerging, and the results of this case study highlight some of the considerations to be addressed in future investigations.

Generalization in the Phonological Domain

In this preliminary case study, we describe an intervention for PD-LD in which a morphophonologically complex consonant sequence was trained in word-final position. The study participant, "Max," was a 5-year-old child presenting with low consonant production accuracy, an especially restricted word-final phonetic inventory, and morphosyntactic omission and agreement errors, including zero marking of word-final grammatical morphemes. He received 14 weeks of treatment targeting a three-element consonant cluster, /–lps/, in word-final position. 3S morphology was embedded in the target by training morphologically rich treatment verb phrases, such as *he helps*.

Max's improvement posttreatment in the phonological domain follows the expectations of a complexity-based approach, which predict expanded phonological knowledge in the form of generalized learning to untrained, simpler linguistic structures (e.g., Gierut & Champion, 2001). After training to produce a complex word-final cluster, Max's phonological system expanded, especially in simpler contexts. Following treatment, he demonstrated the most accuracy improvement in untreated word-initial singletons and clusters. Accuracy improvement in word-final position was large for some sounds and clusters but, overall, was limited and less consistent. Qualitatively, his phonetic inventory following treatment expanded to include four new singletons, five new word-initial clusters, and four new word-final clusters. Given training with a complex word-final target, the asymmetric improvement favoring initial position may seem unexpected. However, given the greater complexity of word-final position relative to initial position (Levelt & van de Vijver, 2004), generalization to simpler structures is expected before generalization of the treatment target or other complex structures is observed (Gierut & Champion, 2001; Taps Richard, Barlow, & Combiths, 2017).

Cross-Domain Generalization

Despite considerable improvement to his phonological system, posttreatment change in Max's morphosyntactic skills was inconsistent. Improvement was notable only in qualitative changes to untreated morphemes in a posttreatment language sample but was unattested in performance on the generalization probes sampling monosegmental regular 3S, past tense, and plural. Learning of the targeted structure, 3S, did not generalize to untreated verbs in generalization probes or a language sample. Furthermore, lack of demonstrated baseline stability for morpheme production in a language sample combined with no significant morphological improvement in the generalization probes suggest that learning did not generalize to the morphological domain. Because 3S morphology was included in the trained cluster, this finding was unexpected but not unprecedented. Cross-domain generalization learning has been attested in several studies; however, there seem to be limits to the directionality of this effect. Studies that simultaneously targeted phonology and morphosyntax in the same session (Tyler et al., 2003) and sequenced a phonology and morphosyntax intervention (Tyler et al., 2002) also demonstrated a similar pattern of posttreatment phonological improvement in the absence of morphosyntactic growth.

However, the limited morphosyntactic improvement in this case study stands in contrast to the outcomes of the children in Tyler and Sandoval (1994), who clearly demonstrated both phonological and morphosyntactic improvement following a cycled treatment that included some training of morphophonemic final clusters. These different treatment outcomes may be attributable to factors related to the participants, the treatment approach, and/or the treatment targets. Participants S1 and S6 (Tyler & Sandoval, 1994) were treated for the phonological process of final consonant cluster reduction by training two-element final clusters in plural-inflected nouns among other cycled targets. During one of his recurring cycles, S1 received direct phonological intervention that included production of plural-inflected nouns with final clusters, similar to Max's intervention presented here. Although S6's cycles also included training of plural-inflected nouns with final clusters, he received a combination of direct phonological intervention and a narrative language intervention that differs more substantially from Max's phonologically based intervention. Consequently, only S1 received treatment comparable in its approach to the treatment described in this study.

Given their pretreatment measures, S1 and S6 (Tyler & Sandoval, 1994) may better fit the profile of children with PD whose morphology errors are ultimately attributable to phonological processes and thus may not be indicative of co-occurring PD-LD. Both participants had language comprehension and expressive vocabulary scores in the normal range prior to treatment. Notably, S1 also demonstrated average MLU for his age. As suggested by the authors, the dramatic improvement of the children's grammatical morpheme production following treatment of word-final phonological processes suggests that phonology was the primary impediment to their production of word-final grammatical morphemes. Thus, participant-internal factors related to impairment profiles may also have contributed to their improvement in the morphological domain.

In addition to potentially different impairment profiles, differences in treatment targets may have impacted the generalizability of learning into the morphosyntactic domain. In Max's intervention, a three-element cluster inflected with 3S was targeted because of its relatively high complexity. Alternatively, in Tyler and Sandoval (1994), two-element clusters with plural *–s* inflection were targeted. Although both targets are complex, two-element clusters are less complex than three-element clusters (Greenberg, 1978), and plural morphology is acquired earlier than 3S and is arguably less complex (Brown, 1973).

Furthermore, there is precedent for limitations to the amount or type of complexity that may contribute to greater generalization learning. For instance, Gierut and Champion (2001) identified less predictable generalization patterns following treatment for PD targeting three-element initial clusters, such as /skw–/, /skɹ–/, or /spl–/. The authors suggested that the unpredictable generalization patterns for these highly complex targets could be explained by each child's pretreatment knowledge of the components of the cluster target. Concisely, functional pretreatment knowledge of two component phonemes resulted in greater generalization learning following treatment. The authors proposed that some knowledge allowed the children to better attend to multiple structural components and generalize learning to a wider variety of other consonant clusters.

Like these initial cluster targets (Gierut & Champion, 2001), Max's word-final treatment target was highly complex. Although he did demonstrate some knowledge of the component singletons /l/ and /p/, he demonstrated poor knowledge of both /l/ and /s/ in word-final position. Thus, it is reasonable to suggest that Max's limited functional knowledge of word-final morphology and of the target's component phonemes impacted his ability to attend to both the phonological and morphological components of his treatment target. Because children presented with novel phonological and grammatical input will, by default, track the input's phonological rather than grammatical characteristics (Plante, Vance, Moody, & Gerken, 2013), it is then less surprising that generalization learning occurred primarily within the phonological domain.

Another potentially impactful difference between Max's treatment and that of S1 and S6 (Tyler & Sandoval, 1994) relates to the variability of the treatment stimuli. Recent findings suggest that treatment of grammatical morphology is more effective when the form and context of the trained structure are highly variable. Presumably, this variability allows children to attend to the stable component of the input (i.e., the grammatical target; Leonard & Deevy, 2017; Plante et al., 2014). Because Tyler and

Sandoval targeted phonological processes, multiple consonant cluster exemplars were employed (at least for S1), and treatment stimuli were not limited to a set of predetermined words. In the case of 3S morphology, regular variants have several phonological forms ([–s], [–z], and [–ɪz]). However, only one variant, [–s], was targeted in Max's intervention, and this form was trained in a limited set of seven treatment verb phrases. It is possible that low variability in the treatment stimuli, in addition to Max's pretreatment knowledge and PD-LD profile, impacted his ability to attend to and generalize learning of the morphological component of his treatment target.

Limitations

The development and execution of this case study closely followed the clinical process as it might occur in school-based practice. A unique profile of phonological and morphosyntactic deficits was identified in a child, for which little treatment efficacy information was currently available. In accordance with the principles of evidence-based practice, a treatment program based on treatment methodologies already shown to be effective in related populations was customized to address the child's unique needs. Although every effort was made to maintain the validity of this study, certain limitations were unavoidable. The study design, including the treatment protocol and baseline/outcome measurements for this study, was developed and executed within the confines of a school environment. Hours spent on treatment and generalization probes were obligated to meet the requirements of the child's IEP and school schedule. However, school breaks between treatment phases and after treatment did permit comparison of change in outcome measures during treatment with the impact of maturation when no treatment was occurring. Also in accordance with the participant's IEP, treatment was provided in a small group, as typically occurs in school settings (Mullen & Schooling, 2010). In this case, one other child of similar age simultaneously received treatment for PD targeting word-initial consonant clusters /skw–/, /spl–/, /stɹ–/, and /θɹ–/. Although Max's secondhand exposure to the other child's initial cluster target was unavoidable, two clinicians were often present to provide directed intervention for each child. When Max attempted a word including the other child's target, no feedback was given, and he was redirected to his own target. Importantly, Max also showed no improvement in the other child's specific word-initial cluster targets over the course of treatment.

Clinical Implications

This exploratory case study suggests that treatment targeting complex three-element clusters in word-final position may be feasible for children with PD-LD, at least for generalization learning within the phonological domain. Perhaps more importantly, these results call for future work to investigate the generalizability of complex exemplars trained in intervention for children with PD-LD. As has been shown separately for treatment of PD and DLD, strategic manipulation of the target stimuli can produce better treatment outcomes (Cummings & Barlow, 2011; Gierut, 2008; Plante et al., 2014; Van Horne et al., 2017), and the generalization principles behind this work should also apply to children with PD-LD. It will be important for future work to clarify the speech-language profiles of children with PD-LD and differentiate overlapping characteristics of PD and DLD from co-occurring PD-LD. Future intervention research should manipulate the phonological complexity of treatment targets as a function of children's pretreatment linguistic knowledge and may choose to determine the impact of target characteristics that have been successful for treatment of DLD, such as variability of the target grammatical form and its context (Plante et al., 2014) or complexity of the stimulus word (Van Horne et al., 2017).

As our ability to identify impairments and disorders improves, comorbidity is becoming the norm rather than the exception, and SLPs faced with the assessment and treatment of children who present with combined speech and language deficits require strategic tools to provide this population with optimal treatment outcomes. Unfortunately, the knowledge our field has accumulated regarding efficacious approaches, targets, and stimuli for PD and DLD is yet to be systematically examined in children with PD-LD. In the meantime, an SLP should ensure that phonology and morphosyntax are thoroughly assessed in children with speech or language deficits so that co-occurring deficits can be either ruled out or targeted. When there is reasonable expectation of overlap or co-occurrence, an SLP should be cautious upon interpreting standardized scores from broad expressive language or phonology measures, as these may be influenced by overlapping characteristics of the nontargeted domain. In these cases, subscores targeting nonoverlapping phonological and morphosyntactic skills may be consulted, or these scores may be supplemented with other probes or performance-based measures. For instance, a phonology probe should include sounds that commonly occur in English inflectional morphology (i.e., /t/, /d/, /s/, /z/) in multiple word positions and in both inflected (e.g., *goes* /goʊz/) and uninflected (e.g., *buzz* /bʌz/) words. Similarly, assessment of tense and agreement morphology should allow for examination of morpheme production in a variety of phonological contexts (e.g., a combination of vowel-final stems, *see* /si/ and *fly* /flaɪ/, and consonant-final stems, *jump* /ʤʌmp/ and *wash* /wɑʃ/). Differences in either domain that are modulated by the other should be considered when interpreting these phonological and morphological assessments.

When co-occurring impairment is identified, a variety of simultaneous or sequenced treatment approaches may be effective; however, treatment targets should be monitored concurrently with progress in any impaired language domains to ensure that generalization is occurring in the targeted domains. Given limited research in this area and the results of this case study, targets may need to be modified if improvement appears to asymmetrically impact one domain over another.

Acknowledgments

The first author was supported by the Lipinsky Fellowship in Language and Communicative Disorders and National Institute on Deafness and Other Communication Disorders Grant T32 (DC007361). Additional funding was provided by National Institute on Deafness and Other Communication Disorders Grant RO3 (DC012141) and a Price Philanthropies research grant awarded to the fourth author. The authors thank Max, his family, and the supportive staff and faculty at his school. They are also grateful to members of the Phonological Typologies Lab and the Child Language Development, Disorders, and Disparities Lab at San Diego State University—in particular, Taylor Kubo, Alyssa Pitts, Doris Yeung, Taylor Trosper, Emily Reynolds, and Elaine Drexler for their contributions to data analyses. They also thank Yvan Rose and Greg Hedlund, creators of Phon.

References

Archibald, L. M., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52*(4), 899–914.

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

Beitchman, J. H., Nair, R., Clegg, M., Ferguson, B., & Patel, P. G. (1986). Prevalence of psychiatric disorders in children with speech and language disorders. *Journal of the American Academy of Child Psychiatry, 25*(4), 528–535.

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., Inglis, A., & Lancee, W. (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(6), 815–825.

Beitchman, J. H., Wilson, B., Brownlie, E. B., Walters, H., & Lancee, W. (1996). Long-term consistency in speech/language profiles: I. Developmental and academic outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(6), 804–814.

Bellon-Harn, M. L., Hoffman, P. R., & Harn, W. E. (2004). Use of cloze and contrast word procedures in repeated storybook reading: Targeting multiple domains. *Journal of Communication Disorders, 37*(1), 53–75.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., & CATALISE-2 Consortium. (2017). Phase 2 of CATALISE: A multinational and multidisciplinary Delphi consensus study of problems with language development: Terminology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 58*(10), 1068–1080.

Bishop, D. V. M., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*(2), 156–173.

Botting, N., & Conti-Ramsden, G. (2003). Characteristics of children with specific language impairment. In L. Verhoeven & H. V. Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 23–38). Psychology Press.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Combiths, P. N., Amberg, R., & Barlow, J. A. (2016). *Phonological analysis software in assessment and treatment of speech-sound disorders*. Paper presented at the Convention of the American Speech-Language-Hearing Association, Philadelphia, PA.

Conti-Ramsden, G., St Clair, M. C., Pickles, A., & Durkin, K. (2012). Developmental trajectories of verbal and nonverbal skills in individuals with a history of specific language impairment: From childhood to adolescence. *Journal of Speech, Language, and Hearing Research, 55*(6), 1716–1735.

Cummings, A. E., & Barlow, J. A. (2011). A comparison of word lexicality in the treatment of speech sound disorders. *Clinical Linguistics & Phonetics, 25,* 265–286.

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,* 309–317.

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.

Fey, M. E., Cleave, P. L., Ravida, A. I., Long, S. H., Dejmal, A. E., & Easton, D. L. (1994). Effects of grammar facilitation on the phonological performance of children with speech and language impairments. *Journal of Speech and Hearing Research, 37,* 594–607.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research, 42,* 708–726.

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16,* 6–17.

Gierut, J. A. (2008). Experimental instantiations of implicational universals in phonological acquisition. In D. A. Dinnsen & J. A. Gierut (Eds.), *Optimality theory, phonological acquisition and disorders* (pp. 355–376). London, United Kingdom: Equinox.

Gierut, J. A., & Champion, A. H. (2001). Syllable onsets II: Three-element clusters in phonological treatment. *Journal of Speech, Language, and Hearing Research, 44,* 886–904.

Gierut, J. A., & Morrisette, M. L. (2010). Phonological learning and lexicality of treated stimuli. *Clinical Linguistics & Phonetics, 24,* 122–140.

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), 1464–1481.

Gillon, G. T. (2000). The efficacy of phonological awareness intervention for children with spoken language impairment. *Language, Speech, and Hearing Services in Schools, 31,* 126–141.

Greenberg, J. H. (1978). Some generalizations concerning initial and final consonant clusters. In J. H. Greenberg, C. A. Ferguson, & E. A. Moravcsik (Eds.), *Universals of human language* (Vol. 2: Phonology). Stanford, CA: Stanford University Press.

Haskill, A. M., & Tyler, A. A. (2007). A comparison of linguistic profiles in subgroups of children with specific language impairment. *American Journal of Speech-Language Pathology, 16*(3), 209–221.

Hodson, B. W., Nonomura, C. W., & Zappia, M. (1989). Phonological disorders: Impact on academic performance? *Seminars in Speech and Language, 10*(3), 252–259.

Hodson, B. W., & Paden, E. P. (1991). *Targeting intelligible speech: A phonological approach to remediation* (2nd ed.). Austin, TX: Pro-Ed.

Hoffman, P. R., Norris, J. A., & Monjure, J. (1990). Comparison of process targeting and whole language treatments for phonologically delayed preschool children. *Language, Speech, and Hearing Services in Schools, 21,* 102–109.

Hoffman, P. R., Norris, J. A., & Monjure, J. (1996). Effects of narrative intervention on a preschooler's syntactic and phonological development. *National Student Speech Language Hearing Association Journal, 23,* 5–13.

Ingram, D. (1976). *Phonological disability in children*. New York, NY: Elsevier.

Ingram, D. (1989). *First language acquisition*. Cambridge, United Kingdom: Cambridge University Press.

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 Disorders, 35*(2), 165–188.

Leonard, L. B. (2014). *Children with specific language impairment* (2nd ed.). Cambridge, MA: MIT Press.

Leonard, L. B., & Deevy, P. (2017). The changing view of input in the treatment of children with grammatical deficits. *American Journal of Speech-Language Pathology, 26,* 1030–1041.

Levelt, C. C., & van de Vijver, R. (2004). Syllable types in cross-linguistic and developmental grammars. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in phonological acquisition* (pp. 204–218). Cambridge, United Kingdom: Cambridge University Press.

Lewis, B. A., Freebairn, L., Tag, J., Ciesla, A. A., Iyengar, S. K., Stein, C. M., & Taylor, H. G. (2015). Adolescent outcomes of children with early speech sound disorders with and without language impairment. *American Journal of Speech-Language Pathology, 24*(2), 150–163.

Lewis, B. A., Freebairn, L. A., & Taylor, H. G. (2000a). Academic outcomes in children with histories of speech sound disorders. *Journal of Communication Disorders, 33*(1), 11–30.

Lewis, B. A., Freebairn, L. A., & Taylor, H. G. (2000b). Follow-up of children with early expressive phonology disorders. *Journal of Learning Disabilities, 33*(5), 433–444.

Macrae, T., & Tyler, A. A. (2014). Speech abilities in preschool children with speech sound disorder with and without co-occurring language impairment. *Language, Speech, and Hearing Services in Schools, 45*(4), 302–313.

Matheny, N., & Panagos, J. M. (1978). Comparing the effects of articulation and syntax programs on syntax and articulation improvement. *Language, Speech, and Hearing Services in Schools, 9,* 57–61.

Miller, J., & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT), *English & Spanish* [Computer software] (Version 8). Madison: Language Analysis Lab, University of Wisconsin.

Montgomery, J., Dunaway, C., & Taps, J. (2005). Interview with Claudia Dunaway and Jennifer Taps. *Communication Disorders Quarterly, 27*(1), 55–58.

Morrisette, M. L., Farris, A. W., & Gierut, J. A. (2006). Applications of learnability theory to clinical phonology. *Advances in Speech-Language Pathology, 8,* 207–219.

Mullen, R., & Schooling, T. (2010). The National Outcomes Measurement System for pediatric speech-language pathology. *Language, Speech, and Hearing Services in Schools, 41*(1), 44–60.

Munson, B., Schellinger, S. K., & Urberg Carlson, K. (2012). Measuring speech-sound learning using visual analog scaling. *SIG 1 Perspectives on Language Learning and Education, 19*(1), 19–30.

Panagos, J. M. (1974). Persistence of the open syllable reinterpreted as a symptom of language disorder. *Journal of Speech and Hearing Disorders, 39*(1), 23–31.

Paul, R., & Shriberg, L. D. (1982). Associations between phonology and syntax in speech-delayed children. *Journal of Speech and Hearing Research, 25,* 536–547.

Pennington, B. F., & Bishop, D. V. (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*(1), 283–306.

Plante, E., Ogilvie, T., Vance, R., Aguilar, J. M., Dailey, N. S., Meyers, C., . . . Burton, R. (2014). Variability in the language input to children enhances learning in a treatment context. *American Journal of Speech-Language Pathology, 23*(4), 530–545.

Plante, E., Vance, R., Moody, A., & Gerken, L. (2013). What influences children's conceptualizations of language input? *Journal of Speech, Language, and Hearing Research, 56*(5), 1613–1624.

Raitano, N. A., Pennington, B. F., Tunick, R. A., Boada, R., & Shriberg, L. D. (2004). Pre-literacy skills of subgroups of children with speech sound disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 45*(4), 821–835.

Rice, M., & Wexler, K. (2001). *Rice/Wexler Test of Early Grammatical Impairment*. San Antonio, TX: The Psychological Corporation.

Rose, Y., & Hedlund, G. (2017). Phon (Version 2.2) [Computer software]. Retrieved from https://www.phon.ca/phontrac

Rvachew, S., Gaines, B. R., Cloutier, G., & Blanchet, N. (2005). Productive morphology skills of children with speech delay. *Journal of Speech-Language Pathology and Audiology, 29,* 83–89.

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

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders II: A conceptual framework for management. *Journal of Speech and Hearing Disorders, 47,* 242–256.

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.

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

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow transcription. *Clinical Linguistics & Phonetics, 5,* 225–279.

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research, 42,* 1461–1481.

Storkel, H. L. (2018a). The complexity approach to phonological treatment: How to select treatment targets. *Language, Speech, and Hearing Services in Schools, 49*(3), 463–481.

Storkel, H. L. (2018b). Implementing evidence-based practice: Selecting treatment words to boost phonological learning. *Language, Speech, and Hearing Services in Schools, 49*(3), 482–496.

Tambyraja, S. R., & Dunkle, J. T. (2014). *Target selection in speech therapy: Is a non-developmental approach more efficient than a developmental approach?* (Vol. 8). Bloomington, MN: Pearson.

Taps Richard, J. (2012). *In-depth Phonological Assessment (IPA)*. San Diego, CA: SLPath.

Taps Richard, J., Barlow, J. A., & Combiths, P. N. (2017). *Applying phonological complexity in the schools: Insights from 32 case studies*. Paper presented at the American Speech-Language-Hearing Convention, Los Angeles, CA.

Thompson, C. K., Ballard, K. J., & Shapiro, L. P. (1998). The role of syntactic complexity in training wh-movement structures in agrammatic aphasia: Optimal order for promoting generalization. *Journal of the International Neuropsychological Society, 4,* 661–674.

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,* 591–607.

Tyler, A. A., Lewis, K. E., Haskill, A., & Tolbert, L. C. (2002). Efficacy and cross-domain effects of a phonology and a morphosyntax intervention. *Language, Speech, and Hearing Services in Schools, 33,* 52–66.

Tyler, A. A., Lewis, K. E., Haskill, A., & Tolbert, L. C. (2003). Outcomes of different speech and language goal attack strategies. *Journal of Speech, Language, and Hearing Research, 46*(5), 1077–1094.

Tyler, A. A., & Sandoval, K. T. (1994). Preschoolers with phonological and language disorders: Treating different linguistic domains. *Language, Speech, and Hearing Services in Schools, 25,* 215–234.

van Daal, J., & van Balkom, H. (2004). Subtypes of severe speech and language impairments: Psychometric evidence from 4-year-old children in the Netherlands. *Journal of Speech, Language, and Hearing Research, 47,* 1411–1423.

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), 2569–2588.

Appendix (p. 1 of 3)

Word Lists

Singletons probe			
cherry	foot	keys	page
chip	toes	gate	peaches
vanilla	thumb	fence	sandwiches
jello	mustache	roof	cages
cereal	beard	off	pages
zero	thick	bathtub	bridges
knife	cage	bath	hop
jar	shell	sink	dance
gum	ocean	chair	jump
chew	sub	vase	hug
yoyo	treasure	chef	hello
think	beach	hat	wave
map	bees	yes	shave
game	beehive	turn	zip
run	fire	mother	dive
leg	jungle	father	lick
yoga	thunder	brother	crash
breathe	edge	valentine	push
race	web	love	scratch
bike	night	necklace	wash
ride	sun	young	sick
hurt	water	mad	cough
ouch	pool	funny	shake
pencil	tube	big	hopping
math	vine	jacket	dancing
read	duck	girl	jumping
long	dog	smooth	hugging
thankyou	fur	television	waving
children	pig	mud	shaving
eighth	cow	muddy	zipping
these	shark	noise	diving
voice	four	noisy	licking
nose	zebra	bridge	crashing
mouth	head	peach	pushing
teeth	lion	badge	scratching
tooth	zoo	book	washing
shaking	coughing		

Appendix (p. 2 of 3)

Word Lists

Clusters probe

snake	swing	cloud	sprouting
slug	splashing	view	green
spider	few	snowflake	smoke
dragon	swimming	truck	plug
school	frisbee	drive	splice
flag	slide	blue	straight
crayons	player	stop	screwdriver
glue	scream	broom	clock
glasses	thrilling	skate	squeezing
black	sprint	quiet	twelve
square	queen	cry	strike
French	throne	stairs	bracelet
fries	princess	music	
shredded	twins	smile	
grapes	tree	cute	
pretzel	flower	pure	
spoon	beautiful	shrub	

Word-final morphophonology probe

board	milks	slippers	rabbits
gulp	pats	sprinkles	behind
pulp	pours	drums	six
scalp	starts	milk	playground
search	stuffs	trunk	stripes
drinks	walks	desk	corn
elks	goes	planets	moths
gulps	Babs	vest	oink
humps	laps	grapes	thinks
skirts	prompt	cats	helps
filled	silk	third	drinks
jumped	friend	horse	melts
missed	church	earmuffs	yelps
cried	inchworm	present	dunked
cries	crabs	blocks	called
drops	scarf	iceberg	steered
drums	gold	park	helped
grabs	airport	games	fanned
jumps	mask	fork	chinked
planned			
yelped			
bird			
klutz			
next			
lynx			
sculpt			
jinx			
sixth			
twelfth			
calx			
mulct			
alps			
whilst			
waltz			
mumps			
glimpse			
prompt			
thousandth			
instinct			
distinct			
length			
infarct			
warmth			
quartz			
corpse			
horst			

Appendix (p. 3 of 3)

Word Lists

Treatment stimuli
he helps
he gulps
he whelps
he pulps
he malps
he telps
he kelps

Chapter 5, in full, is a reprint of material as it appears in Combiths, P., Barlow, J.A., Richard, J.T., & Pruitt-Lord, S.L. (2019). Treatment targets for co-occurring speech-language impairment: A case study. *Perspectives of the ASHA Special Interest Groups, 4*(2), 240–256. The dissertation author was the primary investigator and author of this paper.

CHAPTER 6:

Phonological Complexity in Intervention for Spanish-Speaking Children with Speech Sound Disorder

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, USA;* [b]*Department of Cognitive Science, University of California, San Diego, USA*

*Correspondence concerning this article should be addressed to Philip Combiths,

pcombiths@sdsu.edu,

School of Speech, Language, and Hearing Sciences,

San Diego State University, San Diego, CA 92182, USA

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 a multiple-baseline single-case design across participants to examine the efficacy of intervention in Spanish using different complex targets (i.e. /gɾ/, /bɾ/, /l/). For each of 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 implicational relationships.

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.

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/), and the status of these sequences as “true” branching onset clusters in Spanish is debated. 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 also may 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.

The Current Study

In this initial study, we employ a multiple-baseline (MBL) single-case experimental design to examine the effect of speech intervention targeting relatively complex phonological structures in Spanish, including complex clusters (e.g. /gɾ/) 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:

(1) Is phonological learning stimulated by intervention targeting relatively complex phonological structures?
(2) 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 was approved by San Diego State University's institutional review board (IRB), under protocol number HS-2019-0021. 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 6-1. 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.

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 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 6-1, and inclusionary and exclusionary criteria are specified in table 6-2.

Table 6-1. 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

Note. 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 6-2. Inclusionary and exclusionary criteria.

Inclusionary criteria
<u>*Language profile*</u>
Mexican-US Spanish as first language
Living at home with a native Spanish-speaking caregiver
<u>*Phonological disorder*</u>
≥ 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

Note. 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.

Assessment Procedures

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

Table 6-3. 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

Note. 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).

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 6-4. 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 for experimental control and allowed isolation of an intervention effect by mitigating the confounds of time, maturation, and general variability in performance across probes.

Table 6-4. 0%-at-baseline sound structures monitored for generalization.

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

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).

<u>Experimental Design</u>

This study uses a form of MBL, across-participants design (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 each individual's response to intervention. Third, participants' baseline stability within and across conditions provides control 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. /r/). 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 /gɾ/ 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).

[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.

Phonological targets were embedded in six words used throughout the intervention. These included three real Spanish words (e.g. *grupo* “group”, *grano* “grain”, *grave* “serious/bad”), and three nonwords following the phonotactic restrictions for permissible words in Spanish (e.g. *graki*, *gruka*, *grema*). Nonwords were included as they have been shown to more readily induce generalization learning 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, independently 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

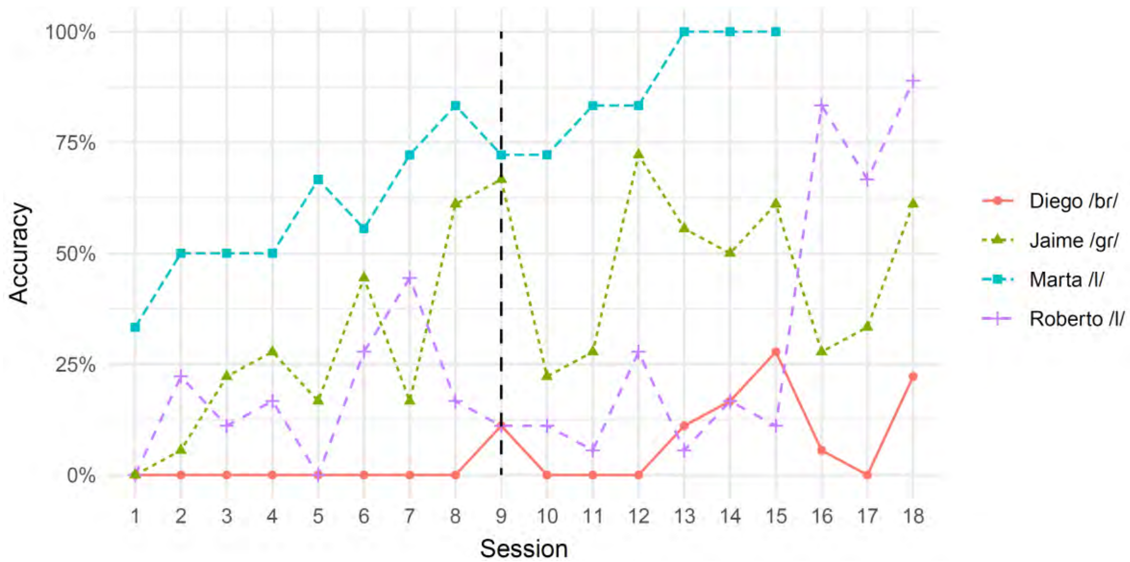
In this section, we provide the results of the study, beginning with a qualitative description of intervention components and each participant's intervention progress in terms of accuracy with the targeted structure in practiced words. Then we describe the dependent variable, which captures broad phonological 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 were used to emphasize the presence of two components in his cluster target, /br/; 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 either 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 were 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 reading included after the midpoint of treatment.

Figure 6-1. Intervention target accuracy.



Note. Dashed line indicates midpoint/phase shift.

Each participant's performance on the target accuracy probe for each intervention session is displayed in figure 6-1. One participant, Marta, 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 exhibited large fluctuations, ending at 61% accuracy with his target. Roberto also fluctuated, falling near or below 10% target accuracy until session 16 where accuracy rose sharply, ending at 89%. Diego did not 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 are displayed in figure 6-2. 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 6-5. 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]

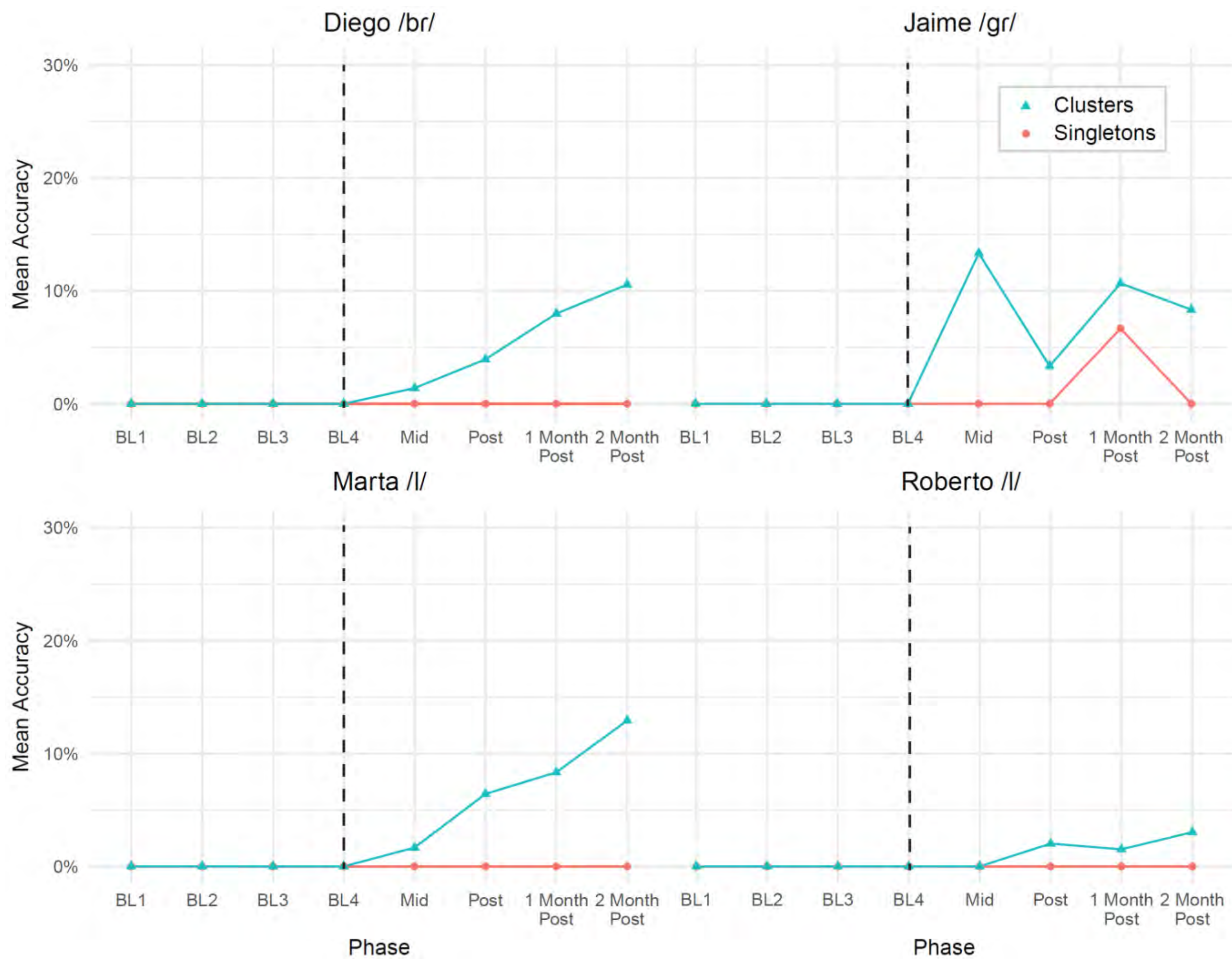
Table 6-5. Accuracy change in monitored clusters.

Name	Target	Pre	Mid	Post	1 Month Post	2 Month Post
Diego	/kl/	0% (0)	0% (0)	20% (0.18)	0% (0)	0% (0)
	/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)
	/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)
	/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)
	/pr/	0% (0)	0% (0)	33% (0.27)	0% (0)	0% (0)

Note. Standard errors in parentheses.

[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 6-2. Composite accuracy of 0%-at-baseline singleton consonants and clusters.



Note. CC = consonant+consonant cluster, CG = consonant+glide cluster, BL = baseline. SD = sonority distance. Dashed line indicates onset of intervention.

Effect Size

To estimate an intervention effect size comparable with similar MBL designs in clinical phonology, standard mean difference was calculated following Gierut 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 are shown in 6-6. 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 6-6 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 6-6 also displays each participant's PCC-R (Shriberg et al., 1997) for comparison with a typical clinical metric.

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

Name	Target	Baseline	Intervention	SMD	PCC-R Pre	PCC-R Post
Diego	/br/	0.3%	3.2%	1.5	57.2%	64.9%
Jaime	/gr/	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%

Note. 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.

Inventories

Each child's phonetic and cluster inventories are shown at pre- and post-intervention timepoints in table 6-7. At the post session, Marta added [ɾ, dw, kw, mj, nj, tj], Roberto added [bj, fj, kj, mj, nj, pj, sj], and Jaime added [v, ɾ, ʧ, dj, fj, nw, pj, sj, tj, 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, ɬw] at their post sessions.

Table 6-7. Phonetic and cluster inventories at pre and post sessions.

Name	Pre	Post
Diego	[p, b, t, d, k, g, m, n, ɲ, f, v, s, ʝ, x, h, ʦ, ʧ, w, β, ɣ, ð, l, ɾ, j]	[p, b, t, d, k, g, m, n, ɲ, f, s, ʝ, x, h, ʧ, w, β, ɣ, ð, l, j]
	[fw, sw, bw, pw, kw, tw]	[fw, sw, bw, pw, kw, tw]
Jaime	[p, b, t, t^h, d, k, g, m, n, ɲ, f, s, ʝ, x, h, ɬ, ʦ, w, β, ɣ, ð, l, j]	[p, b, t, t^h, d, k, g, m, n, ɲ, f, **v**, s, ʝ, x, ɬ, ʦ, **ʧ**, w, β, ɣ, ð, l, **ɾ**, j]
	[ɬw, k^{ə}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	[p, b, t, d, k, g, ʔ, m, n, ɲ, f, v, s, ʝ, x, h, w, β, ɣ, ð, l, j]	[p, b, t, d, k, g, m, n, ɲ, f, s, ʝ, x, h, w, β, ɣ, ð, l, **ɾ**, j]
	[dj, bw]	[**nj**, **mj**, dj, **dw**, **kw**, **tj**]
Roberto	[p, b, t, d, k, g, m, n, ɲ, f, v, s, ʃ, ʝ, x, ʦ, ʧ, w, β, ɣ, ð, l, l̃, ɾ, j]	[p, b, t, d, k, g, m, n, ɲ, f, v, s, ʝ, x, ʧ, w, β, ɣ, ð, l, ɾ, j]
	[dj, tj]	[**nj**, **mj**, **sj**, **fj**, **bj**, **kj**, **pj**]

Note. Singletons or cluster added to the inventory in the post-intervention session are bolded.

<u>Intelligibility in Context Scale and Parent Report</u>

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 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 at 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 /br/. 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 /gr/. 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 was slower than all but Diego to demonstrate learning of the intervention target, despite a singleton target being ostensibly easier to master than a cluster target (e.g. Rvachew & Nowak, 2001). Although somewhat delayed, 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.

<u>Clinical and Theoretical Implications</u>

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/, /pr/) 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 those 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 do 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 which 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 form of MBL, across-participants design. It differed from a canonical MBL 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.

<u>Summary and Future Directions</u>

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/, /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 is 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.

Acknowledgements

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, NIDCD R21 DC01720, and NIDCD T32 DC007361.

References

Almost, D., & Rosenbaum, P. (1998). Effectiveness of speech intervention for phonological disorders: A randomized controlled trial. *Developmental Medicine and Child Neurology, 40*, 319-325.

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.), (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.

Barberena, L. d. S., 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.

Barlow, J. A. (2003). Asymmetries in the acquisition of consonant clusters in Spanish. *Canadian Journal of Linguistics, 48*, 179-210.

Barlow, J. A. (2005). Phonological change and the representation of consonant clusters in Spanish: A case study. *Clinical Linguistics & Phonetics, 19*(8), 659-679.

Barlow, J. A. (2012). *Little PEEP: Shorter Protocol for the Evaluation of English Phonotactics* [PowerPoint Presentation]. Phonological Typologies Lab, San Diego State University.

Barlow, J. A., & Combiths, P. N. (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), 3-3.

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.

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.

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*, 446-472.

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.

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.

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.

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.

Elbert, M., & McReynolds, L. V. (1985). The generalization hypothesis: Final consonant deletion. *Language and Speech, 28*(3), 281-294.

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.

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*, 67-75.

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.

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.

Fogel, A., & Thelen, E. (1987). Development of early expresive and communicative action: Reinterpreting the evidence from a dynamic systems perspective. *Developmental Psychology, 23*(6), 747-761.

Gierut, J. A. (1990). Differential learning of phonological oppositions. *Journal of Speech and Hearing Research, 33*(3), 540-549.

Gierut, J. A. (1991). Homonymy in phonological change. *Clinical Linguistics & Phonetics, 5*(2), 119-137.

Gierut, J. A. (1998a). Natural domains of cyclicity in phonological acquisition. *Clinical Linguistics and Phonetics, 12*(6), 481-499.

Gierut, J. A. (1998b). Treatment efficacy: Functional phonological disorders in children. *Journal of Speech, Language, and Hearing Research, 41*, S85-S100.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research, 42*(3), 708-726.

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16*(1), 6-17.

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. In *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.

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.

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.

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), 1464-1464.

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.

Gierut, J. A., & Neumann, H. J. (1992). Teaching and learning /θ/: A non-confound. *Clinical Linguistics & Phonetics, 6*(3), 191-200.

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.

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*.

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.

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*, 924-943.

McLeod, S. (2020). Intelligibility in Context Scale: Cross-linguistic use, validity, and reliability. *Speech, Language and Hearing, 23*(1), 9-16.

McLeod, S., & Crowe, K. (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*, 1546-1571.

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.

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.

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*, 214-229.

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.

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.

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*, 45-57.

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.

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.

Restrepo, M. A. (1998). Identifiers of predominantly Spanish-speaking children with language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1398-1411.

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].* In

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.

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

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]. San Diego.

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.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders II: A conceptual framework for management. *Journal of Speech and Hearing Disorders, 47*, 242-256.

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.

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow transcription. *Clinical Linguistics & Phonetics, 5*(3), 225-279.

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.

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.

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.

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.

Thelen, E. (1995). Motor development: A new synthesis. *American Psychologist, 51*(2), 79-95.

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*, 591-607.

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), 2569-2569.

Watts, E., & Rose, Y. (2020). Markedness and implicational relationships in phonological development: A cross-linguistic investigation. *International journal of speech-language pathology*, 1-14.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2012). *Preschool Language Scales Fifth Edition Spanish*. Pearson.

Chapter 6, in full, is a reprint of the manuscript, Combiths, P., Pruitt-Lord, S., Escobedo, A., & Barlow J.A. (submitted). Phonological complexity in intervention for Spanish-speaking children with speech sound disorder. The dissertation author was the primary investigator and author of this paper.

CHAPTER 7:

General Discussion

The goal of this dissertation is to advance clinical approaches to assessment and treatment for children with PD, leveraging our understanding of phonology, phonological analysis, and phonological development to improve the efficiency of these approaches. Given health disparities in speech-language pathology (see Holt & Ellis, 2016) that can negatively affect those who do not happen to fall within the most studied populations (i.e., White, monolingual speakers of majority varieties of English), an additional goal of this work is to advance assessment and treatment in ways considerate of the clinical and linguistic diversity of children with phonological impairments in service of a more equitable and inclusive evidence base. To set the stage for this work, Chapter 1 provided an overview of the current state of assessment and treatment for PD, with a lens toward areas for improvement of clinical methodologies that could advance the efficiency and equity of these approaches in the extant research.

In Chapters 2 and 3, potentially more efficient approaches to independent phonological assessment were explored. First, the AutoPATT plugin was introduced, and the accuracy and validity of its automated inventory analyses (i.e., phonetic, phonemic, and cluster inventories) were examined in comparison to the same independent assessment measures derived manually. This study found AutoPATT to be significantly more accurate than manual analyses. Improved accuracy, alone, indicated that these automated assessment tools can advance the efficiency of linguistically motivated, independent analyses, which are otherwise cumbersome and time consuming. Their efficiency may be further increased when the potential time-saving aspects of automation are also considered. In a second study, phonemic inventories generated with AutoPATT were compared to a percent accuracy measure to determine how accuracy is correlated with functional phonemic usage of consonants in children with PD. By examining the relationship between phonemic status and accuracy in the single-word productions of 275 English-speaking preschool and early school-age children with PD, we found that up to 90% of

English consonants could be classified correctly as either in or out of a given child's phonemic inventory based on an accuracy cutoff between 20–30%. Given the relative ease and familiarity with accuracy analyses among practicing clinicians, this relationship between two measures could offer a more efficient alternative to a qualitative, linguistically motivated analysis. Given the findings of both chapters, we can improve the efficiency of thorough phonological assessment procedures. Further, with more accessible tools for descriptive, independent analysis, we can offer clinicians more ecologically valid assessment options for diverse populations with PD (Castilla-Earls et al., 2020; McLeod & Baker, 2014), more accurate identification of impairment (Fabiano-Smith, 2019; Fabiano-Smith & Goldstein, 2010), and better informed decisions about treatment targets (Barlow et al., 2011; Morrisette et al., 2006; Storkel, 2018).

In Chapter 4, we examined the influences of surrounding phonological context on explicit marking of two tense and agreement morphemes in the English productions of Spanish-English bilingual preschool children. For both past tense /-d/ and third-person singular /-z/ morphemes, sonority of the immediately preceding segment in the verb stem impacted marking rate, although for third-person singular /-z/, this effect was dependent upon the probability of syllabification into the following word. An explanation for this differential pattern of surrounding phonological influence involved syllable- and word-level phonotactics, their interaction with syllabification into the onset of the following word, and the influence of Spanish phonotactic constraints on the English productions of these bilingual children. This work highlighted the interaction between phonology and morphology in word-final tense and agreement marking in English, an aspect of grammar that has been well established as a clinical marker for DLD in English-speaking children (Bedore & Leonard, 1998; Rice & Wexler, 1996). However, tense and agreement marking rate, alone, is problematic as an indicator of DLD for multilingual children or those who use a non-majority variety of English (Gutiérrez-Clellen et al., 2008; Iglesias & Rojas, 2012; Paradis & Crago, 2000; Paradis et al., 2008; Pruitt & Oetting, 2009; Pruitt et al., 2011). Consequently, the results of this study are relevant to the use of tense and agreement marking *patterns*, rather than marking *rate* as diagnostic indicators of DLD in Spanish-English bilingual children. Finally, the influence of phonology on variable production of English word-final grammatical morphology highlights the possibility of further exploration of inflected verb-final position as a potential nexus for treatment simultaneously targeting phonological and morphological deficits in English.

In Chapter 5, a case study was presented in which a morphophonologically complex structure, /-lps/ in English verbs inflected for third-person singular present tense, was targeted in treatment for an English-speaking child with co-occurring PD and DLD. This study was motivated by the need to examine the role of linguistic complexity in across-class and across-domain generalized learning in the relatively prevalent but understudied population of children with PD who also demonstrate deficits in other areas of language (i.e., morphosyntax; Paul & Shriberg, 1982; Rvachew et al., 2005). Following expectations of language learning and complexity theory (e.g., Gierut, 2007), the child in this study demonstrated across-class phonological growth in areas of related or lesser complexity than the treatment target. However, across-domain generalization to grammatical morphology was not observed. As a case study, this work highlighted the need for systematic manipulation of word-final treatment targets to determine if greater treatment efficiency can be achieved in the form of cross-domain change from treatment with a single morphophonologically complex target.

In Chapter 6, four Spanish-speaking bilingual children with PD received treatment targeting a Spanish consonant cluster (/gr/, /br/) or singleton (/l/) to observe the effect of treatment, which differed in the structural complexity of the target, on broad phonological generalization patterns. Given the notable absence of research comparing targets for phonological treatment provided in Spanish, this work sought to improve the equity of treatment for PD by addressing access to intervention in the first language of a bilingual demographic that is both large and understudied (Kohnert & Medina, 2009). In this study, children who received treatment targeting a complex consonant cluster demonstrated more across-class phonological growth attributable to treatment than children who received treatment with a singleton target. This study thus extended comparable findings for treatment in English (Gierut, 1999; Gierut & Champion, 2001), providing evidence for the efficiency of treatment targeting complex clusters in Spanish for children with PD. By observing patterns of across-class phonological generalization in a language with considerable phonological differences from English, the language that has been almost exclusively examined in this regard (e.g., Gierut, 2015), we also gained insight into the limitations of cross-linguistic predictions based exclusively on implicational relationships (Watts & Rose, 2020) and the structural complexity of word-initial consonant+glide sequences (e.g., /fw/) in Spanish development. An additional finding in this study was that phonological variability at baseline may be indicative of readiness for change (e.g., Rvachew & Bernhardt, 2010; Smith & Thelen, 2003) and, consequently, the extent to which broad phonological generalization may be expected to occur following complexity-based treatment.

Taken together, these studies provided converging evidence for the importance of phonological structure in the assessment and treatment of PD, with implications for efficiency and equitable access to intervention. Phonological complexity was implicated in the findings of all the studies described here, excepting the validation study for AutoPATT. In particular, the role of branching syllable structure in consonant clusters was shown to impact explicit marking of word-final tense and agreement morphemes in the English productions of typically developing Spanish-English bilingual children (Chapter 4). Complex branching clusters have been linked to broad, across-class phonological growth when incorporated into treatment targets for PD (Gierut, 1999; Gierut & Champion, 2001), and this work extended these findings to a child with a profile of co-occurring phonological and morphosyntactic impairment (Chapter 5) and Spanish-speaking bilingual children with PD (Chapter 6).

The work described in this dissertation was primarily exploratory in nature and has also generated new directions for continued research in this area. Regarding independent phonological assessment tools, the existence of software, like AutoPATT, is unlikely to be impactful without a better understanding of how clinicians with diverse caseloads are likely to use such a tool successfully. Future work should draw on principles of implementation science, which have gained increasing traction in speech-language pathology (Baker et al., 2018; Burns et al., 2020; Cunningham & Oram Cardy, 2020; Douglas & Burshnic, 2019; McGill & McLeod, 2020; McLeod, 2020; Olswang & Prelock, 2015; Sugden et al., 2018; Watts Pappas et al., 2016), to determine the most effective strategies for clinical implementation of this and other tools designed to improve efficiency or provide better access to assessment and treatment.

The case study and single-case design studies examining generalization outcomes following phonological treatment in Chapters 5 and 6 highlighted the considerable variability that children may demonstrate in response to this type of complexity-based intervention. This variability must be reconciled with the theoretical motivation for this work and the specifics of the observed outcomes. As discussed in both studies, implicational relationships attested across languages are thought to be the primary linguistic motivation for observed patterns of generalization following treatment with a relatively complex structure (e.g., a consonant cluster; Gierut, 2007). However, Watts and Rose (2020) highlight the need for caution when ascribing a causal or implicational relationship to ostensibly unrelated structures in phonological development (e.g., clusters imply affricates) based on typological observations outside of the context of acquisition. This is especially problematic when the vast majority of translational work examining treatment outcomes based on these implicational predictions (which are derived from cross-linguistic data) has only been conducted in English (i.e., *not* cross-linguistic data). In their study, which examined longitudinal acquisition data for young speakers of French, German, English and Portuguese, some, but not all of the implicational relationships that have been associated with complexity-based target selection were attested in their cross-linguistic acquisition data. Relevant to the work described here, the prediction that consonant clusters imply singletons was robustly attested across their data; however, more specific predictions, such as clusters imply affricates or liquid onset clusters imply a liquid in the coda, were either unattested or found inconsistently. Similarly, in the generalization patterns following treatment for the Spanish-speaking bilingual children with PD, outcomes were generally consistent with implicational predictions; however, there were also examples of unpredicted growth, notably the emergence of branching onset clusters in a child trained with a less complex singleton target. In short, this work highlighted that constraints within a child's phonological system tend to follow patterns attested in fully formed languages, but the idiosyncrasies of as-yet-unknown child-internal factors and a developing phonological system that interacts with and adapts to other developing cognitive and motor skills create an ever-changing dynamic system that may not always be fully predictable (Smith & Thelen, 2003).

Consequently, it will be important for future work to critically examine the multiple influences on generalization outcomes to better isolate the active ingredients of target selection and maximize the efficiency of treatment. It may be the case that the most relevant considerations include certain aspects of phonological complexity and child-internal phonological knowledge, but these are likely to interact idiosyncratically with each child's other developing systems, such that the optimal treatment target considerations would vary across children—and these are not limited only to phonological factors. One avenue to proceed with this work is to challenge our existing conceptualizations of linguistically optimal treatment targets by continuing to study treatment in the context of languages with phonological structures that differ from those that have been studied to date (i.e., English and Spanish). It is through cross-linguistic examination that we are best able to understand phonological phenomenon, and this is true for fully formed languages as well as language systems in development—with or without impairment.

As this work continues, it will benefit greatly from a cross-theoretical, transdisciplinary approach to better illuminate the linguistic and non-linguistic components that actively contribute to efficient treatment for PD, and—crucially—how these components interact with each other. Combined with outcome data from phonological interventions conducted in a wider variety of languages and a greater diversity of impairment profiles, we can seek to better understand the nuances of assessment and treatment for PD as it occurs in the diverse array of children who would stand to benefit from this work. Thus, we have in front of us a feasible pathway to advance clinical service provision for PD that is more efficient, and a more equitable approach to this research will not only improve the accessibility of treatment but also provide a better cross-linguistic lens through which we can better understand phonological impairments and how best to mitigate their impact.

References

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.

Barlow, J. A., Taps, J., & Storkel, H. L. (2011). Phonological Assessment and Treatment Target (PATT) Selection (Spanish) [Assessment protocol].

Bedore, L. M., & Leonard, L. B. (1998). Specific language impairment and grammatical morphology: A discriminant function analysis. *Journal of Speech, Language, and Hearing Research, 41*, 1185-1192.

Burns, C. L., Wishart, L. R., Kularatna, S., & Ward, E. C. (2020). Knowing the costs of change: An introduction to health economic analyses and considerations for their use in implementation research. *Speech, Language and Hearing, 23*(1), 30-36.

Castilla-Earls, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., & Peña, E. (2020). Beyond scores: Using converging evidence to determine speech and language services eligibility for dual language learners. *American Journal of Speech-Language Pathology, 29*(3), 1116-1132.

Cunningham, B. J., & Oram Cardy, J. (2020). Using implementation science to engage stakeholders and improve outcome measurement in a preschool speech-language service system. *Speech, Language and Hearing, 23*(1), 17-24.

Douglas, N. F., & Burshnic, V. L. (2019). Implementation science: Tackling the research to practice gap in communication sciences and disorders. *Perspectives of the ASHA Special Interest Groups, 4*(1), 3-7.

Fabiano-Smith, L. (2019). Standardized tests and the diagnosis of speech sound disorders. *Perspectives of the ASHA Special Interest Groups, 4*(1), 58-66.

Fabiano-Smith, L., & Goldstein, B. (2010). Assessment and intervention in bilingual children. In S. Levey (Ed.), *Language Development, Disorders, and Differences*. Sage.

Gierut, J. A. (1999). Syllable onsets: Clusters and adjuncts in acquisition. *Journal of Speech, Language, and Hearing Research, 42*(3), 708-726.

Gierut, J. A. (2007). Phonological complexity and language learnability. *American Journal of Speech-Language Pathology, 16*(1), 6-17.

Gierut, J. A. (2015). Participant eligibility and demographics. In *Learnability Project Working Papers*. Indiana University.

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.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G., & Wagner, C. (2008). Bilingual children with language impairment: A comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29*(1), 3-19.

Holt, Y., & Ellis, C. (2016). Racial disparities: An old problem that needs a new approach in speech-language pathology-issues in pediatric populations. *ECHO: Journal of the National Black Association for Speech-Language and Hearing, 11*(1), 100-106.

Iglesias, A., & Rojas, R. (2012). Bilingual language development of English language learners: Modeling the growth of two languages. In *Bilingual language development and disorders in Spanish-English speakers* (2nd ed., pp.3-30). Brooks.

Kohnert, K., & Medina, A. (2009). Bilingual children and communication disorders: A 30-year research retrospective. *Seminars in Speech and Language, 30*(4), 219-233.

McGill, N., & McLeod, S. (2020). Waiting list management in speech-language pathology: Translating research to practice. *Speech, Language and Hearing, 23*(1), 2-8.

McLeod, S. (2020). Intelligibility in Context Scale: Cross-linguistic use, validity, and reliability. *Speech, Language and Hearing, 23*(1), 9-16.

McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics & Phonetics, 28*(7-8), 508-531.

Morrisette, M. L., Farris, A. W., & Gierut, J. A. (2006). Applications of learnability theory to clinical phonology. *Advances in Speech-Language Pathology, 8*(3), 207-219.

Olswang, L. B., & Prelock, P. A. (2015). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research, 58*(6), 1-14.

Paradis, J., & Crago, M. (2000). Tense and temporality: A comparison between children learning a second language and children with SLI. *Journal of Speech, Language, and Hearing Research, 43*(4), 834-834.

Paradis, J., Rice, M. L., Crago, M., & Marquis, J. (2008). The scquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics, 29*(4), 689-722.

Paul, R., & Shriberg, L. D. (1982). Associations between phonology and syntax in speech-delayed children. *Journal of Speech and Hearing Research, 25*, 536-547.

Pruitt, S. L., & Oetting, J. B. (2009). Past tense marking by African American English-speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 52*, 2-15.

Pruitt, S. L., Oetting, J. B., & Hegarty, M. (2011). Passive participle marking by African American English-speaking children reared in poverty. *Journal of Speech, Language, and Hearing Research, 54*(2), 598-607.

Rice, M. L., & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1239-1257.

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.

Rvachew, S., Gaines, B. R., Cloutier, G., & Blanchet, N. (2005). Productive morphology skills of children with speech delay. *Journal of Speech-Language Pathology and Audiology, 29*, 83-89.

Smith, L. B., & Thelen, E. (2003). Development as a dynamic system. *Trends in Cognitive Sciences, 7*(8), 343-348.

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.

Sugden, E., Baker, E., Munro, N., Williams, A. L., & Trivette, C. M. (2018). Service delivery and intervention intensity for phonology-based speech sound disorders. *International Journal of Language and Communication Disorders, 53*(4), 718-734.

Watts, E., & Rose, Y. (2020). Markedness and implicational relationships in phonological development: A cross-linguistic investigation. *International journal of speech-language pathology*, 1-14.

Watts Pappas, N., McAllister, L., & McLeod, S. (2016). Parental beliefs and experiences regarding involvement in intervention for their child with speech sound disorder. *Child Language Teaching and Therapy, 32*(2), 223-239.