Journal of Child Language (2022), **49**, 80–113
doi:10.1017/S0305000920000768

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Phonetic discrimination, phonological awareness, and pre-literacy skills in Spanish–English dual language preschoolers

Sara A. SMITH[1], Sibylla LEON GUERRERO[2], Sarah SURRAIN[2], and Gigi LUK[3]

[1]University of South Florida, USA, [2]Harvard Graduate School of Education, USA, and [3]McGill University, Canada
Corresponding author: Sara A. Smith University of South Florida, College of Education.
E-mail:Sarasmith3@usf.edu

(Received 27 August 2019; revised 31 July 2020; accepted 15 November 2020;
first published online 11 February 2021)

Abstract
The current study explores variation in phonemic representation among Spanish–English dual language learners (DLLs, n = 60) who were dominant in English or in Spanish. Children were given a phonetic discrimination task with speech sounds that: 1) occur in English and Spanish, 2) are exclusive to English, and 3) are exclusive to Russian, during Fall (age $m = 57$ months) and Spring (age $m = 62$ months, n = 42). In Fall, English-dominant DLLs discriminated more accurately than Spanish-dominant DLLs between English-Spanish phones and English-exclusive phones. Both groups discriminated Russian phones at or close to chance. In Spring, however, groups no longer differed in discriminating English-exclusive phones and both groups discriminated Russian phones above chance. Additionally, joint English-Spanish and English-exclusive phonetic discrimination predicted children's phonological awareness in both groups. Results demonstrate plasticity in early childhood through diverse language exposure and suggest that phonemic representation begins to emerge driven by lexical restructuring.

Introduction

During the first year of life, infants go through a period of PERCEPTUAL NARROWING in which they hone emerging speech perception to focus on the linguistic sounds salient in the ambient environment (Burns, Yoshida, Hill & Werker, 2007; Werker & Tees, 1984; Kuhl et al., 2006). Infants learn the sounds of their language, the distributions of these sounds, how they combine, and relevant prosodic information (rhythm, stress, syllables; Curtin & Zamuner, 2014). As children expand their lexicon, sound representations are further refined and restructured (Carroll, 2004; Metsala & Walley, 1998; Ainsworth, Welbourne, Woollams & Hesketh, 2019). For children growing up in a bilingual environment, however, naturalistic and intensive exposure to multiple languages can occur before, during, and after this period of perceptual narrowing. Bilingual development is additionally shaped by onset age of acquisition and linguistic features in the second language (L2) in addition to developmental and linguistic features in the first language (L1). Among bilingual children, variations in language input are particularly influential for L1 and L2 acquisition trajectories and outcomes (see Hoff, 2018, for a review). Prior research indicates that domains related to quality and quantity of early language input are vocabulary, morphosyntax (Paradis, Nicoladis, Crago & Genesee, 2011), phonological perception (Unsworth, 2016), and phonological working memory (Pierce, Genesee, Delcenserie & Morgan, 2017). Specific influential features suggest that total exposure in L1 and L2, variety and richness of input, languages present in home and school environments, and timing of exposure (simultaneous, sequential) matter for bilingual children's language outcomes (see Unsworth, 2016, for a review). Common to many bilingual children in a dominant language community, exposure to L2 may come from schooling where the instructional language is not the same as their L1 or home language. The current study seeks to examine variation in English phonemic perception as it relates to L1 and L2 dominance to elucidate phonemic development in bilingual children.

Given the different definitions of phonological structures that are of interest in the current study, we present the operational definitions adopted in the paper. PHONEME refers to abstract categories that define minimal contrasts between lexical items in a language (Ainsworth et al., 2019). PHONES are the distinct speech sounds that carry meaning, regardless of a contrast or specific language and PHONETIC refers to all speech sounds and their physical, physiological, and acoustic characteristics, independent of a specific language (Ainsworth et al., 2019). Phonemes are relative to their specific language, i.e., a Spanish /p/ and an English /p/ may not have identical distributions of acoustic features, but may overlap (Hayes-Harb, 2007). PHONEMIC REPRESENTATION refers to the abstract phonemic category storage of the sound structure in words. PHONEMIC PERCEPTION develops sequentially aligning with the size of linguistic structures, starting from larger units (i.e., syllables) to smaller (i.e., onset rime) as the child grows older (Metsala & Walley, 1998). PHONOLOGICAL REPRESENTATIONS are the mental representations of the sounds that occur in words.

PHONEMIC SENSITIVITY refers to perception of speech sounds but does not necessarily include a distinction between speech sounds (Ainsworth, Welbourne & Hesketh, 2016; Ainsworth et al., 2019). PHONEME DISCRIMINATION refers to the ability to recognize a distinction between, or distinguish between, contrasting sounds: for example, being able to recognize phonemes as different. PHONEMIC DISCRIMINATION refers to distinguishing between acoustically similar sounds (differing frequency, duration, intensity) that distinguish meaning in the language (i.e., the phonemes and the underlying representation stored in the mental lexicon). PHONETIC DISCRIMINATION refers to distinguishing between sounds that are different, even if the two sounds are not differentiated in a specific language. IMPLICIT sensitivity refers to children's unconscious knowledge of the sounds of words while EXPLICIT AWARENESS of sounds segments refers to being consciously aware of one's knowledge of the sounds and structure of words (Ainsworth et al., 2016; 2019). PHONOLOGICAL AWARENESS describes the ability to reflect on and manipulate the phonological segments within the representations (i.e., metacognitive, explicit knowledge of these phonological elements; Ainsworth et al., 2019). In the present study, we examine English- or Spanish-dominant DLL preschoolers and their explicit phonetic discrimination ability longitudinally after exposure to English instruction.

Language acquisition in bilingual children

Children growing up in a bilingual environment are exposed to more diverse speech sounds that are socially meaningful to them (for a review, see Hoff, 2018) and as such, bilingual speakers develop differently from monolinguals of either language (Sadeghi & Everatt, 2017). While the general sequence of achieving milestones is comparable, bilingual experience shifts this developmental timeline based on phonetic features of the two languages as well as quality and quantity of exposure. The overarching consensus now supports differentiated lexical, syntactic, and phonological systems for developing bilingual children, as outlined by Genesee's (1989) Differentiated Language System Hypothesis. While some features of language such as lexical and syntax development are generally, though not exhaustively, well researched, the extent to which the phonological representations of two languages are differentiated in development is still inconclusive (for a review, see Quay & Montanari, 2016).

Bilingualism is associated with an extended period of "receptiveness" or plasticity; bilingual infants and children remain sensitive to characteristics of speech sounds after monolinguals have ceased to be (see Birdsong, 2018 for a review; Ferjan Ramirez, Ramírez, Clarke, Taulu & Kuhl, 2017). Bilingual infants can differentiate between their multiple languages: however, acquiring overlapping contrasts (i.e., sounds that are close in both languages' sound systems) may present a challenge. This is attributed to the intrinsic challenge of acquiring more phonemes overall and being exposed to two potentially "imperfectly overlapping" phonetic distributions (Fennell, Tsui & Hudon, 2016, p. 46). Research among Spanish-Catalan bilingual infants describe a unique U-shaped developmental pattern for discriminating phoneme contrasts that exist in Catalan but not in Spanish; bilingual infants successfully discriminate Catalan- exclusive vowels at 4 months, fail to discriminate at 8 months, before successfully discriminating again after 12 months (Bosch & Sebastián-Gallés, 2003). This temporary "struggle" has been attributed to the close nature of some sounds in different languages, resulting in overlapping distributions. Contrasts without clear bimodal distributions present overlapping data; this must be resolved (i.e., identify exact end points for contrasts) to discriminate between the two sounds (Maye, Werker & Gerken, 2002).

Oral language experience predominantly drives lexical restructuring and the emergence of phonemes (Ventura, Kolinsky, Fernandes, Querido & Morais, 2007). As children acquire a larger verbal repertoire, their phonological representation is likely derived from the words being added to the child's expanding lexicon, in turn influencing how new words are learned; there appears to be a reciprocal relationship between phonological and lexical knowledge development (Curtin & Zamuner, 2014). Children reorganize phonetic segments, form new phonological representations, and tweak existing representations (Anthony, Solari, Williams, Schoger, Zhang, Branum-Martin & Francis, 2009; although for findings of no relationship between vocabulary size and detection of mispronounced words; see Swingly & Aslin, 2000).

Development of phonological representations is highly influenced by input; children developing two phonological systems show variation in degree of phonetic discrimination in each language related to quantity of input (for a review, see Byers-Heinlein & Fennell, 2014; Garcia-Sierra, Ramírez-Esparza & Kuhl, 2016). Garcia-Sierra, Ramírez-Esparza, and Kuhl (2016) examined neural responses to

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

native and non-native speech sound and found that quantity of input impacted native and non-native phonemic discrimination (as measured through ERP). Results indicated that greater input in a specific language is associated with enhanced neural commitment to that language. Thus, we hypothesize that oral language experiences among young bilinguals, including features of home and school language input, should influence the development of individual sound representations in the words being acquired, i.e., children with more Spanish input will have greater neural commitment to the speech sounds of Spanish and children with more English input will show greater neural commitment to English.

By the age of 3, bilingual children display different phonological development compared to typically developing monolinguals of either language (Gildersleeve, Davis & Stubbe, 1996; Goldstein & Washington, 2001). For example, Spanish–English bilinguals show phonological patterns in their two languages not present among monolinguals of either language (e.g., final consonant devoicing in Spanish, initial consonant deletion in English), coupled with error patterns found in both (cluster reduction, stopping, and gliding; Gildersleeve et al., 1996). Furthermore, Spanish–English bilingual children also make errors (e.g., substitution patterns) not observed in monolinguals of either language (Goldstein & Washington, 2001). These errors made uniquely by bilingual children could be interpreted as evidence of ongoing phonological system development. Specifically, sounds with overlapping distributions, the frequency of phonemes in each language, and restructuring related to an expanding bilingual lexicon might result in changing or imprecise boundaries for speech sounds. While young bilinguals initially produce more errors than monolinguals, this difference generally decreases over time (Gildersleeve-Neumann & Davis, 1998), further evidence that increased input, i.e., more "data", supports reorganization.

Theories of phonological representation development

The field has yet to reach a consensus regarding phonological development. Multiple theoretical frameworks, some competing, are still under exploration. If and when the sound structures of words are stored as abstract phonemic categories remains equivocal. Debate exists over whether: 1) children have adult-like detailed phonological representations and access these only once they have the metacognitive and letter knowledge to do so (e.g., the ACCESSIBILITY ACCOUNT, Liberman, Shankweiler & Liberman, 1989); or 2) words are not initially stored as adult-like phonological representations but increasing adult-like categories emerge as a result of lexical restructuring (e.g., the EMERGENT ACCOUNT, Metsala & Walley, 1998). For a more complete summary of differing theoretical accounts and the predictions each makes about the role of vocabulary and letter sound knowledge, see Ainsworth and colleagues (2019).

Evidence indicates that children likely refine their sensitivity to the sounds of words as their lexicon grows, becoming aware of progressively smaller phonetic segments, from entire word to onset–rime, and ultimately resulting in precise phonological representations (Ainsworth et al., 2019; Metsala & Walley, 1998). The LEXICAL RESTRUCTURING HYPOTHESIS, for example, posits that infants first acquire holistic, word-level phonological representations (overall sound pattern of the word), followed by phonetic segments, and continue to develop and restructure these phonetic segments during the first years of life (see Metsala & Walley, 1998 for a review). This process is distinct from children's acquisition of explicit phonological awareness, which appears to be more related to literacy instruction (Ainsworth et al., 2019).

That is to say, as children learn new words, they appear to go from storing the overall sound of the word to storing the individual sounds within the word.

Alternative theories emphasize the role of phonetic perception as facilitatory for learning new words. Examples include PRIMIR (Werker & Curtin, 2005), which posits rich representations are present from infancy (unlike the vague representations posited by LEXICAL RESTRUCTURING HYPOTHESIS) and phonemic categories are later sharpened through exposure to orthographic representations. Kuhl's native language magnet model, expanded (NLM-e; 2008), describes 4 phases of phonetic development: 1) infants discriminate all phonetic units; 2) detection of relevant phonetic cues is ENHANCED while detection of irrelevant patterns is REDUCED; 3) honed language-specific speech perception abilities support further phonetic refinement; 4) finally, the infant has stable representations. Viewed through the lens of the NLM-e, infants exposed to two languages spends a longer period of time in "phase 2" as it takes longer to get a sufficient amount of input experiences from each language (Kuhl et al., 2008). The purpose of our study is to extend beyond infancy and examine preschoolers' phonetic discrimination.

Distribution-based perceptual learning models propose that the process of learning salient sounds relies on and is refined by the statistical properties of the target language (Werker, Pons, Dietrich, Kajikawa, Fais & Amano, 2007). Infants detect and organize speech sounds based on the distributional properties of their ambient language. Research indicates that developmental patterns of phonetic discrimination align with differences in the statistical distributions of input (Maye, Werker & Gerken, 2002). For example, the frequency of a particular sound is key for acquiring the capacity to discriminate between contrasts, and the order in which phoneme categories are acquired corresponds to the relative frequency of the sounds (Anderson, Morgan & White, 2003). Statistical distribution properties appear to be relevant for a number of features of spoken language discrimination, including the phonotactic sequence (Mattys, Jusczyk, Luce & Morgan, 1999) and word unit (Shi, Werker & Morgan, 1999). Additionally, variations in language-specific phoneme features appear to be relevant, including acoustic distance (Eilers, Wilson & Moore, 1977), salience, and acoustic strength of the contrast (Polka, Colantonio & Sundara, 2001).

Until they begin learning about letters, children may not be conscious of their knowledge regarding the structure of words. The ability to segment and MANIPULATE phonemes is typically acquired when children receive formal reading instruction (Sadeghi & Everatt, 2017). Explicit knowledge of phonological elements allows the child to reflect on and manipulate phonological segments, i.e., have PHONOLOGICAL AWARENESS, which is a predictor of alphabetic literacy (for a meta-analysis, see Bus & van IJzendoorn, 1999). In languages with an alphabetic writing system, reading acquisition relies on children's ability to grasp the "alphabetic principle" – knowledge that there is a correspondence between the phonemes that form a spoken word and the sequence of letters that form a printed word.

Carroll (2004) suggested that letter-sound knowledge is a requirement for sensitivity to speech sounds, such as syllable similarity, rime, and phoneme isolation: however, recent work suggests that letter-sound knowledge may not be needed for IMPLICIT sensitivity to segments. Ainsworth and colleagues (2019) demonstrated that letter-sound knowledge was not requisite for phonemic sensitivity, when measured through tasks that did not require spelling knowledge. Children with little letter–sound knowledge could blend phonemes but could not explicitly describe the sounds within a word. Letter-sound knowledge did, however, play a role in explicit segmenting tasks (Ainsworth et al., 2019). Implicit sensitivity to word segments has a demonstrated relationship with vocabulary growth; however, implicit sensitivity to segments does not show relationships with orthography while explicit segmental tasks are related to letter-sound knowledge (Ainsworth et al., 2019). In this sense, phonological awareness likely requires letter–sound knowledge.

Collectively, the literature suggests that factors related to language input such as the sounds of words, sound distributions, auditory features, and oral language comprehension and production all contribute to the development of phonological representations (Curtin & Zamuner, 2014). Lexical restructuring and emergence of phonemes within the lexicon appears to be highly, although not exclusively, related to oral language experiences (Ventura et al., 2007). Additionally, implicit phonemic sensitivity has the potential to emerge even in the absence of letter-sound knowledge (Ainsworth et al., 2019).

Bilingualism in the U.S.

In the U.S., children learning two languages are often referred to as dual language learners (DLLs) within the education context. Approximately 60% of U.S. DLLs, or roughly 6.9 million, speak Spanish (Park, Zong & Batalova, 2018). There is enormous diversity among U.S. Spanish–English speakers related to home and community language exposure (Park et al., 2018; Hoff, 2018). Previous findings demonstrate differences related to Spanish and English quantity and quality of input (Hoff, 2018) and English and Spanish output (Kim, Lambert & Burts, 2018). Many DLLs from Spanish-speaking households receive schooling in English during early childhood – for example, in Head Start programs (U.S. federally subsidized preschool) – and continue on to English-medium elementary schools. Twenty-two percent of the approximately one million children enrolled in Head Start in 2017-18 heard Spanish regularly at home (~230,000) (Office of Head Start, 2018). To date, however, little research has investigated Spanish-speaking children's phonological plasticity in response to L2 exposure in formal education.

The development of L2 (English) phoneme categories among DLLs is critical to understanding phonological processing, a pre-literacy skill important for subsequent reading development. There is general consensus that the ability to process phonological information is strongly predictive of reading development for alphabetic languages (Adams, 1990; Branum-Martin, Tao, Garnaat, Bunta & Francis, 2012; Snowling, 2001; Torgesen, 2004). Phonological processing is essential for fluent reading, as decoding unfamiliar words (i.e., applying the alphabetic principle) requires children to use phonological awareness and alphabet knowledge (Byrne, 2014). Children who struggle with applying the alphabetic principle tend to also struggle with reading (Byrne, 2014). A recent systematic review identified only a limited number of dual language phonological awareness interventions for DLLs (Soto, Olszweski & Goldstein, 2019). Increased knowledge regarding phonemic development among U.S. Spanish–English DLLs is relevant to developing effective phonological awareness interventions, which have the potential to reduce risk for reading difficulties (Storch & Whitehurst, 2002).

Current study

The unique language landscape in the U.S. has provided an opportunity to investigate how systematic exposure to a second language through schooling shapes DLLs'

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

phonemic development. Hammer and Miccio (2006) followed simultaneous and sequential Spanish–English bilingual Head Start preschoolers in English-only preschool for two years and found that, despite initial differences in preschool, the two groups were comparable in their Spanish and English phonological awareness, letter identification, and letter-sound knowledge in kindergarten.

However, despite a growing population of U.S. DLLs and increased research interest, there is still a paucity of research among DLL children of kindergarten age and younger (Barac, Bialystok, Castro & Sanchez, 2014; Hammer, Hoff, Uchikoshi, Gillanders, Castro & Sandilos, 2014; Kim, Lambert & Burts, 2018), particularly in how the perceptual system responds to L2 exposure as early learning begins. One obstacle in developing this research is assessment. Young bilinguals may not have the vocabulary or conceptual understanding needed to complete conventional picture, rhyming, blending or segmentation tasks often used with school-aged children. In the current study, we created a novel, preschooler-friendly task, adapted from prior tasks ("Who said it better?"; Carroll & Snowling, 2004; Fowlert, Swainson & Scarborough, 2004; Rvachew, Ohberg, Grawburg & Heyding, 2003), designed to assess phonetic discrimination with minimal requirements for overt production or semantic knowledge. Specifically, we investigated whether discrimination of phonetic segments differed for DLLs who predominantly spoke English at home and school, consistent with their English-dominant schooling (English-dominant DLLs), and those whose dominant language was the L1 (Spanish), and who therefore were developing different languages in different contexts, i.e., Spanish at home and English at school (Spanish-dominant DLLs). In addition, we aimed to examine how DLLs' phonetic discrimination develops in preschool and how this development is related to the development of phonological awareness and pre-literacy skills in the school language, i.e., English between two time points in the preschool year. Our research questions were:

1: How is language dominance (Spanish, English) among DLLs related to children's phonetic discrimination of differing types of phones?
2: Do relationships among language dominance, phone type, and phonetic discrimination change from Fall to Spring as children accumulate exposure to English in school (4–6 months)?
3: Is Fall phonetic discrimination related to Spring early literacy skills such as phonological awareness and print knowledge? Further, is this relationship mediated by Spring phonetic discrimination skills?

Methods

Participants

The current study included 60 Spanish–English DLL preschoolers (32 females) attending federally subsidized Head Start programs largely serving low to middle income families. Spanish–English bilingual researchers tested children twice at their schools, once in Fall and once 4–6 months later in Spring. Participating preschools maintained a student-teacher ratio between 7:1 and 10:1 with at least one Spanish–English bilingual teacher in each classroom. All programs designated time for literacy activities such as letter-name instruction and story time, as well as a variety of other activities such as free play, outdoor play, and arts and crafts. Head Start programs receive common curriculum guidance and learning goals at both federal and state levels. In the preliteracy domain, children's learning goals at 60 months of age include recognizing and producing simple rhymes and beginning sounds in words, as well as beginning to identify a few letter names and sounds (Enriquez, 2015). In the programs participating in this study, English was the primary language of instruction, and all children were exposed to Spanish via peer and teacher interactions. Participating children had been enrolled in preschool or daycare and consequently exposed to English instruction for a minimum of 6 months up to 5 years. None of the participants were exposed to Russian in their home, school or community environment.

Preschoolers were 47–65 months old at first testing in the Fall (time 1, $m = 56.9$ months, $sd = 4.4$ months; time 2, $m = 61.8$ months; $sd = 4.8$). All children were recruited through the Head Start programs that they attended at the time. Parents who provided consent were interviewed at the Head Start centers during drop-off in the morning or pick-up in the afternoon using a demographic and home language background questionnaire (adapted from Luk & Bialystok, 2013). Then, bilingual Spanish–English research assistants visited the centers to conduct one-on-one assessments with the children.

Children's language dominance was triangulated using three measures, two indirect (parent & teacher reports) and one direct (Simon Says task). Before the child assessments, parent and teacher reports of children's language dominance were collected via the background questionnaire. A bilingual research assistant then administered a Simon Says task (Leon Guerrero, Smith & Luk, 2016). In this screener, children were given six verbal commands in Spanish (e.g., "Simón dice, tócate la cabeza") and six commands in English (e.g., "Simon says, touch your head"). One point was given for each correct motor response, with a maximum score of six for each language. All children responded correctly to at least four out of six commands in English. All children whose parents and teachers reported Spanish dominance were able to complete two or more Spanish commands ($m = 5.1$), while the majority of children (20/28) reported as English dominant could complete fewer than two Spanish commands ($m = 1.4$). Children's scores on the Spanish Simon Says task were highly correlated with parental reports of children's daily use of Spanish at home (Kendall's tau = 5.75, $p < .001$). Indeed, no child's Simon Says task results were in conflict with parent or teacher reports of language dominance. Spanish-dominant DLLs were then assessed in Spanish (except for assessments of English language) by Spanish–English bilingual researchers while the English-dominant DLLs were assessed in English. Each child's language of testing remained the same in both Fall and Spring.

Table 1 reports sample mean home language usage and Simon Says performance in English and Spanish. Spanish-dominant DLLs had significantly higher Spanish usage at home (66% of daily speech in Spanish) than English-dominant DLLs (15.8%). Between Fall (time 1) and Spring (time 2), 18 children left the program or the area. The final sample includes 60 children at time 1 (32 Spanish-dominant, 28 English-dominant) and 42 children at time 2 (27 Spanish-dominant, 15 English-dominant).

In addition to language variables, families were asked to report the highest level of maternal education (seven categories from elementary to graduate school) and the combined annual family income (six categories from <20K to >100K) via the background questionnaire. Due to the paucity of responses in the lowest and highest categories, we combined maternal education responses into three categories: (1) secondary/high school or less, (2) college, (3) graduate/professional school. Two-thirds of families reported secondary/high school (as opposed to college or graduate school) as the highest maternal level of education, with a similar proportion

Table 1. Language and home environment for the total sample and the Spanish-dominant and English dominant subgroups.

Variable	Spanish dominant DLL (n = 32) m (sd)	English dominant DLL (n = 28) m (sd)
% of time Spanish is spoken at home (n = 58)	66.0 (18.1)	15.8 (16.0)
% of time Spanish is spoken at home with mother (n = 57)	87.5 (17.7)	18.1 (19.5)
% of time Spanish is spoken at home with father (n = 52)	69.7 (26.4)	20.1 (21.4)
% of time Spanish is spoken at home with siblings (n = 52)	42.5 (36.9)	12.5 (14.4)
% of time Spanish is on TV when child is watching (n = 58)	20.7 (24.3)	15.9 (18.7)
% of time Spanish is used at home when reading stories to the child (n = 56)	52.6 (35.5)	12.3 (15.7)
Number of adults residing in home (n = 56)	2.8 (1.7)	1.7 (0.9)
Number of children residing in home (n = 56)	2.8 (1.1)	2.7 (1.3)
Simon says English (n = 60)	5.6 (0.7)	5.9 (0.4)
Simon says Spanish (n = 60)	5.1 (1.0)	1.4 (1.9)

reporting a combined family income of less than $40,000 per year (Table 2). A substantial portion of the sample declined to report income (n = 14) while fewer (n = 4) declined to report maternal education. On average, families of Spanish-dominant children reported lower maternal education (Fisher's exact $p = .01$) than English-dominant ones. There was no significant difference in family income between the two language groups (Fisher's exact $p = .07$).

Measures

All children were given standardized and experimental assessments of language and literacy. An additional executive functions task was also given to the children and has been reported elsewhere (Leon Guerrero et al., 2016).

Standardized measures administered in the child's dominant language

The Kaufman Brief Intelligence Test, 2nd Edition, Matrices subtest (KBIT-2; Kaufman & Kaufman, 2004) assesses children's nonverbal reasoning ability by asking children to point to an image that completes a presented series or analogy. The Matrices subtest was administered in Spanish for Spanish-dominant children and in English for English-dominant ones. Split-half reliability coefficients for the KBIT-2 range from .80 to mid- .90. Fall age-corrected standard scores ($M = 100$, $SD = 15$) were used in the subsequent analysis.

The Clinical Evaluation of Language Fundamentals, Preschool, 2nd Edition (CELF-P2; Wiig, Secord & Semel, 2004) is a standardized measure of language skills for preschool-aged children that has been co-normed in both English and Spanish with a

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

Table 2. Spanish dominant and English dominant DLLs background information

	Total sample (n = 60)	Spanish dominant DLLs (n = 32)	English dominant DLLs (n = 28)
Mother's education level			
High school (12th grade/ GED or less)	40	29	11
Undergraduate (BA or less)	15	2	13
Graduate/Professional School	1	0	1
Not reported	4	1	3
Family income			
Low ($39,000 or less)	40	24	16
Middle ($40,000-$79,000)	6	1	5
High ($80,000 or greater)	0	0	0
Not reported	14	7	7
School attendance			
6 to 11 months	15	8	7
1 year	5	3	2
2 years	19	11	8
3 years	13	7	5
4 years	5	2	3
5 years	1	0	1
Not reported	3	1	2

U.S. sample. Three subtests of oral language ability were administered in the children's dominant language, either in English or Spanish: 1) Concepts and Following Directions: a child points to a series of pictures in an order dictated by orally presented directions; 2) Sentence Structure: a child selects a picture representing a spoken sentence among three distractors; 3) Word Classes: a child is asked to select two related pictures out of three total. Average split-half reliability scores of these subtests range from .80 to .87. The CELF-P2 was administered in both Fall and Spring. Age-corrected scaled scores ($M = 10$, $SD = 3$) were used in the subsequent analysis.

Standardized measures administered in English only

The KBIT-2, Verbal Knowledge & Riddles subtests (Kaufman & Kaufman, 2004) assess children's English word knowledge and reasoning by requiring children to select one out of four pictures that best describes an English word or completes an English riddle spoken by the researcher. We used a Fall age-corrected composite English verbal reasoning standard score ($M = 100$, $SD = 15$) in the subsequent analysis as a measure of English verbal abilities. We did not consider these English verbal subtests as intelligence measures due to potential language bias for Spanish-speaking children.

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

The CELF-P2, Phonological Awareness subtest (Wiig et al., 2004) is an English language assessment of a child's awareness of English words, syllables and phonemes. We administered four tasks in this subtest in Spring, including asking the child to blend two words to form a compound word; blend syllables to form a word; listen to words and judge whether the words rhyme; and produce a word that rhymes with a given word. These tasks yielded a raw score representing total correct responses (maximum = 16).

The English Print Knowledge subtest (Test of Preschool Early Literacy, TOPEL, Ambrose, Fey & Eisenberg, 2012) was administered in Spring. This subtest measures alphabetic knowledge and familiarity with written language conventions and form in English. The test consists of 36 items that require the child to identify individual letters, letter names, match letters to sounds and vice versa, and identify written words. Age-corrected standard scores ($M = 100$, $SD = 15$) were used in subsequent analysis.

Experimental measure

The phonetic discrimination task was designed to test whether children were able to differentiate two single syllable "words" that differ by onset sound. In this task, children were asked to determine if two different "elves" said the same word. These single-syllable words rhyme, but half of them had different onset sounds. Examples of stimulus pair with different onset sounds are "kip"/"gip", "bing"/"ving", "zop"/ "tzop". Three categories of onset sounds were presented: FAMILIAR phones that are salient in English and also occur in Spanish (e.g., l/k); ENGLISH-EXCLUSIVE phones that are salient in English but not in Spanish (e.g., th/d; Schnitzer, 1997); and UNFAMILIAR phones that are phonemes in an unfamiliar language (Russian) but not in English or Spanish (e.g., ш/ш'). As described in the introduction, phonemes are relative to their specific language – for example, a phoneme may have different distributions of acoustic features; a Spanish phoneme /g/ and an English phoneme /g/ may not have the same distribution of acoustic features but may overlap. The FAMILIAR speech sounds presented in the task have a so-called "analogue" phoneme contrast in both English and Spanish; i.e., /g/ occurs in both Spanish and English and is discriminated from /m/. The /g/ sound in English and Spanish may not be identical due to different distributions of acoustic features, but they are considered analogues for the purpose of the current study. The task presented: phones that are phonemes in English exclusively; phones that are phonemes in English and Spanish; phones that are phonemes in Russian. However, because it is unknown if participants have phoneme representations, the task items will be referred to as phones throughout. Thirty-two trials were presented in which 17 of them had the elves speaking the same syllables (congruent trials) and 15 of them spoke syllables with different onset consents (incongruent trials). The contrasts are presented in Table 3, and each incongruent pair was presented once.

All the single-syllable items were consonant-vowel-consonant (CVC) pseudowords that follow English phonemic patterns (ending in "ng" or "p") to cue children to English. Spanish often has a consonant-vowel-consonant-vowel (CVCV) structure, although single-syllable words ending in consonants do exist in Spanish (e.g., *mar*, *sol, pez*). English structure was presented to prime learners for English. Prior research indicates that bilinguals perceive speech differently depending on which language they believe they are listening to (Gonzales, Byers-Heinlein & Lotto, 2019); we alerted participants that they were listening to English. All test items had "p" or

Table 3. Incongruent pairs of phone contrasts

Familiar pair (IPA)	English-exclusive pair (IPA)	Unfamiliar pair (IPA) [Cyrillic]
l/k (l/k)	b/v (b/v)	j/zj (dʒ/ʐ) [j/ж]
s/z (s/z)	sh/ch (ʃ/tʃ)	z/tz (zʲ/t͡s) [з/ц]
d/g (d/g)	d/th (d/ð)	ch/tz (t͡ɕ/t͡s) [ч/ц]
g/m (g/m)	t/th (t/ð)	sh/sh (ʃ/ʂ) [ш/ш’]
	j/ch (dʒ/tʃ)	

“ng” in the final syllable position, ending sounds that do not occur in Spanish (Jiménez González & García, 1995). In Spanish, only five consonant sounds can end a syllable (d, n, s, r, l). Vowel sounds presented in the target items are present in both English and Spanish (e.g., o, i, a, e; Jiménez González & García, 1995). Russian trials presented CVC pseudowords consisting of a Russian onset consonant and vowel-consonant combination.

Children were introduced to the task through a video introduction in either Spanish or English. Children then completed a series of training trials with verbal and visual feedback until it was clear the task was understood. The training trials presented a CVC example that ended in “n”, then “p”. All trial words were recorded using the same female speaker, a Russian-English bilingual. Introductory videos and training were recorded using a female Spanish–English bilingual. The task was presented on a touch screen laptop with accompanying child-sized over-the-ear headphones to ensure the test items were delivered. A splitter was used to deliver the stimuli to the researchers simultaneously, who listened to the task with the child to ensure working sound.

In each trial, two elves appeared on the screen. Each elf spoke a word in turn, followed by a prompt for the child to respond. Children could touch either of two “buttons” on a touch screen: one button presented two identical squares, indicating that the words were the same, or a second button with a square and a triangle, indicating that the words were different. This child-friendly paradigm did not require productive speech from the child and was presented in PsychoPy (Peirce, 2009). Trial-level accuracy and mean accuracy rates from this task were used as data in the study analyses.

Procedures

All research procedures were approved by the Harvard University ethics board. Children were assessed after obtaining center, teacher, and parental consent, as well as child assent. Children were tested in one 60-minute session with breaks as needed. The KBIT-2 Matrices, CELF-P2 Concepts & Following Directions, Sentence Structure, and Word Classes, as well as the experimental discrimination task were administered in the child’s dominant language. The KBIT-2 Verbal Knowledge and Riddles, the CELF-P2 Phonological Awareness, and the TOPEL Print Knowledge subtests were administered in English. All measures were given to children twice separated by 4 to 6 months, except for the Print Knowledge and the CELF-P2 Phonological Awareness subtests. These measures were only administered in Spring (i.e., after 4–6 months of instruction) as they assessed English preliteracy skills and knowledge that reflected learning goals targeted in preliteracy instruction.

Analysis

Prior to analysis, data inspection revealed that the distribution of children's phonetic discrimination task results violated assumptions of normality, and, in regression models, of sphericity and homoscedasticity. We thus utilized non-parametric tests in examining basic descriptive statistics, correlations and first-level group comparisons. Specifically, we employed Wilcoxon signed rank tests and BCa (bias corrected and adjusted) bootstrapped 95% confidence intervals (CIs) to test differences in sample means; Kendall's test of association (tau) to examine pairwise (zero order) correlations among numeric variables of interest; and one-way Kruskal-Wallis ANOVAs to examine descriptive associations between numeric and categorical variables of interest.

For the first two research questions examining the relationship of language dominance and phone type with children's phonetic discrimination task, our outcome variable was trial-level task accuracy. We modeled dichotomous trial responses, i.e., a score of "1" for a correct answer and "0" for an incorrect answer, using multilevel logistic regression models with subject random intercepts through R (R Core Team, 2018) package *lme4* (Bates, Mächler, Bolker & Walker, 2015). The model taxonomy for each research question included: 1) a baseline model with only the main effect of phone type, including child age (in months) as a control variable in order to account for potential effects of maturation on phone task performance. In addition, maternal education was included in all baseline models as a control proxy for family socioeconomic status. Participants with missing maternal education data ($n = 4$ in Fall; $n = 3$ in Spring) were removed listwise in models that included this variable. Maternal education was not a significant predictor in the baseline models for RQ 1 and RQ2, and did not substantively alter the magnitude or sign of the remaining model coefficients. Therefore, we removed the maternal education variable from subsequent models in the interest of parsimony and conserving model degrees of freedom. 2) Building from the baseline model, the second model in each taxonomy included language group as the main predictor of substantive interest as well as time of testing for the longitudinal research questions. 3) We then introduced subsequent model(s) that added interactions of the predictors of interest to their main effects. 4) Lastly, we selected the final model by comparing model fit across the taxonomy using ANOVAs of model deviance and the Akaike information criterion (AIC). Where model fit was not significantly different, we selected the most parsimonious model as final. Residual plots (residuals vs. fitted values, quantile, and residuals vs. leverage) for all models were examined to ensure that model assumptions were not violated. Multicollinearity of variables in each model was evaluated using the variance inflation factor (VIF). Parameters in all models displayed a VIF less than 2, indicating very low multicollinearity of model predictors.

We employed a similar model-building procedure for the third research question examining performance on the two pre-literacy measures: TOPEL Print Knowledge standard scores and CELF-P2 Phonological Awareness raw scores as our dependent variables in taxonomies of generalized linear regression models. Because performance on pre-literacy assessments has been associated with IQ (e.g., Lonigan, Schatschneider & Westberg, 2008; Milburn, Lonigan & Phillips, 2019) and with maternal education (e.g., Bus, van Ijzendoorn & Pellegrini, 1995; Curenton & Justice, 2008; Rowe, Denmark, Harden & Stapleton, 2016) in prior literature, we included KBIT2 nonverbal reasoning and maternal education scores as control variables in a generalized linear regression baseline model in addition to age. When these variable coefficients were not statistically significant in the baseline model and did not substantively alter the magnitude or sign of remaining model coefficients, we removed the variable from all subsequent models in the interest of model parsimony and power. Control variables with significant coefficients in the baseline model were retained throughout the remainder of the model taxonomy.

In order to construct the baseline model, we determined the best-fitting distribution to use in each model taxonomy through visual inspection, substantive alignment, and likelihood ratio tests. First, because the TOPEL Print Knowledge subtest norms have discreet, positive bounds and our sample data was strongly negatively skewed despite its origins in a standardized, norm-referenced assessment, in the generalized regression model taxonomy predicting this variable we employed a Gaussian distribution truncated at the minimum and maximum values of our sample distribution and employing the standardized test mean ($\mu = 100$) and standard deviation ($\sigma = 15$). This truncated distribution provided the best fit to the sample outcome as measured by likelihood ratio tests. Second, as CELF-P2 Phonological Awareness subtest raw scores are count measures, in this model taxonomy we employed a generalized Poisson distribution, which provided the best fit to the sample outcome as measured by likelihood ratio comparisons. All models for the third research question were fit and compared using the R package family *gamlss* (Rigby & Stasinopoulos, 2005).

Results

Descriptive statistics

Standardized measures

When tested in their dominant language in Fall, Spanish- and English-dominant DLLs performed comparably on standardized measures of nonverbal reasoning (Table 4, Kruskal-Wallis X^2=3.1, $p = .20$). Similarly, when tested in their dominant language with the CELF-P2 in Fall, Spanish- and English-dominant DLLs also did not differ on average in performance on any subtests (Table 4, p >.05 for all Kruskal-Wallis X^2 values). However, mean CELF-P2 scores in both groups were significantly lower than the population mean (all Wilcoxon $p < .05$; BCa bootstrapped 95% CIs did not include 10.0, the population mean) on all subtests except for Spanish-dominant DLLs' CELF-P2 Word Classes, which did not differ from the population mean (Wilcoxon p =.04; BCa bootstrapped 95% CI [8.28, 10.0]). In Spring, both groups showed evidence of 'catching up', or average scores closer to the population mean than in Fall. Across the entire sample, median difference between all Fall and Spring CELF-P2 subtest standard scores was 1.0 point, or one-third of a standard deviation, and the two language groups did not differ significantly in the magnitude of the difference between Fall and Spring scores (p>.05 on all Kruskal-Wallis X^2 values). By Spring, Spanish-dominant DLLs did not differ on average from the population norm on Spanish Sentence Structure and Word Classes, Receptive subtests (both Wilcoxon $p > 0.5$, bootstrapped 95% CIs include 10), while the English-dominant DLLs displayed lower performance in English Sentence Structure and Word Classes in Spring compared to the population norm (all Wilcoxon $p < .05$).

In contrast, when tested on English-only measures, Spanish-dominant DLLs scored lower on average than English-dominant ones, as might be expected from their lower

Table 4. Means and standard deviations of standardized task performance for Spanish dominant and English dominant DLLs in time 1 and time 2.

	Time 1		
	Spanish dominant DLL, (n = 32) m (sd)	English dominant DLL (n = 28) m (sd)	Group difference (p of Kruskal-Wallis χ^2)
KBIT-2 non-verbal ability[a]	92.5 (10.9)	95.6 (13.1)	3.1 (.20)
KBIT-2 verbal ability[a]	75.9 (11.3)	95.9 (13.8)	20.0 (<.001)
CELF-P2 Concepts and Following Directions[b]	7.9 (3.1)	7.7 (2.6)	0.2 (.86)
CELF-P2 Sentence Structure[b]	8.3 (2.2)	8.6 (2.6)	0.3 (.62)
CELF-P2 Word Classes[b]	9.1 (2.6)	8.4 (3.3)	0.7 (.88)
	Time 2		
	Spanish dominant DLL (n = 27) m (sd)	English dominant DLL (n = 14) m (sd)	Group difference (p of Kruskal-Wallis χ^2)
CELF-P2 Concepts and Following Directions[b]	8.4 (2.8)	7.2 (1.6)	1.2 (.18)
CELF-P2 Sentence Structure[b]	9.6 (2.5)	8.1 (2.0)	1.5 (.05)
CELF-P2 Word Classes[b]	10.1 (2.7)	7.7 (2.9)	2.4 (.007)
CELF-P2 Phonological Awareness raw score (max = 16)	5.9 (3.6)	8.5 (5.2)	2.6 (0.18)
TOPEL Print awareness in English[a]	92.7 (12.8)	105.5 (11.3)	12.8 (.004)

[a]Standardized test with population mean of 100 and standard deviation of 15.
[b]Standardized test with population mean of 10 and standard deviation of 3.

levels of English input and use. The two language groups differed significantly on English language measures in Fall English KBIT-2 Verbal Reasoning scores, with English-dominant DLLs on average not different from the population mean (Wilcoxon $p = .16$; BCa bootstrapped 95% CI [90.71, 100.96]) and Spanish-dominant DLLs significantly lower on average than their English-dominant peers (Kruskal-Wallis $X^2 = 20.0$, $p < .001$). Similarly, on the Spring English Print Knowledge (TOPEL), although Spanish-dominant DLLs scored on average within one SD of the population mean, this group performed significantly lower than English-dominant DLLs (Kruskal-Wallis $X^2 = 8.1$, $p = .004$).

Phonetic discrimination

On average in Fall, children were able to discriminate phones in the experimental task with 63% accuracy. Accuracy varied, however, by phone type, with children scoring highest on average with familiar phones (71%), followed by English-exclusive phones (61%) and unfamiliar phones (57%). When disaggregated by language group and phone type, both Spanish- and English-dominant DLLs were able to discriminate familiar (for both groups Wilcoxon $p < .001$) and English-exclusive (for Spanish-dominant DLLs Wilcoxon $p = .01$ and English-dominant DLLs $p < .001$) phones above chance. Additionally, Fall bootstrapped 95% CIs for familiar and English-exclusive phones (see Table 5) did not include the chance level (at 0.5). However, for unfamiliar phones, only the English-dominant group performed significantly above chance in Fall (Wilcoxon $p = .006$). Children's discrimination improved over time with an average increase of 9.7 percentage points from Fall to Spring, and in Spring, both groups performed above chance in all phone categories (all Wilcoxon $p < .05$; no 95% CIs include 0.5) as seen in table 5. Numerically, the largest average difference between Fall to Spring occurred with familiar phones (12.3%) as compared to English-exclusive (11.6%) and unfamiliar phones (6.0%).

Research question 1: phonetic discrimination and language dominance

In research question 1, we asked whether language dominance and phone type predict DLLs' ability to distinguish phones. In the Fall baseline mixed-effects logistic regression models (table 6, Model 1.1) DLLs on average had a 71% probability of correctly discriminating familiar phones ($\beta = .89$, $p < .001$), significantly higher than for English-exclusive phones at 62% ($\beta = -.42$, $p < .001$) or unfamiliar phones at 56% ($\beta = -.65$, $p < .001$). When comparing DLLs by language group in Model 1.2 across all phone types, the odds of English-dominant vs. Spanish-dominant DLLs for correctly identifying phones was 1.38:1 ($\beta = -.32$, $p < .001$).

However, this main effect of language group was moderated by phone type as the final, best-fitting Model 1.3 (Table 6) included significant main effects of phone type and language group as well as the interaction between these terms. English-dominant DLLs were more likely than Spanish-dominant DLLs to be successful on familiar (odds = 1.6:1, $\beta = -.45$, $p = .008$) and English-exclusive (odds = 1.5:1, post-hoc $\beta = -.43$, $p = .02$) phone trials. On average, the probability of correctly discriminating familiar phones was estimated to be 77% for English-dominant DLLs ($\beta = 1.22$, $p < .001$) and 68% for Spanish-dominant DLLs ($\beta = -.45$, $p = .008$), illustrated in Figure 1. Similarly, the probability of correctly discriminating English-exclusive phones was estimated to be 68% for English ($\beta = -.47$, $p < .001$) and 58% for Spanish-dominant DLLs (post-hoc $\beta = -.43$, $p = .02$). However, the two language groups did not differ

Table 5. Means, standard deviations, and BCa bootstrapped 95% confidence intervals of phone contrasts in the phonetic discrimination task for Spanish dominant and English dominant DLLs at time 1 and time 2.

	Fall (Time 1)			
	Spanish dominant DLL (n = 32)		English dominant DLL (n = 28)	
	m (sd)	BCa bootstrap 95% C.I.	m (sd)	BCa bootstrap 95% c.i.
Familiar phone contrasts	0.67 (0.19)	(0.60, 0.73)	0.76 (0.14)	(0.71, 0.81)
English exclusive phone contrasts	0.57 (0.17)	(0.52, 0.63)	0.67 (0.16)	(0.61, 0.73)
Unfamiliar phone contrasts	0.55 (0.17)	(0.49, 0.61)	0.59 (0.15)	(0.53, 0.65)
	Spring (Time 2)			
	Spanish dominant DLL (n = 27)		English dominant DLL (n = 15)	
	m (sd)	BCa bootstrap 95% C.I.	m (sd)	BCa bootstrap 95% c.i.
Familiar phone contrasts	0.77 (0.19)	(0.69, 0.83)	0.85 (0.14)	(0.75, 0.90)
English exclusive phone contrasts	0.70 (0.15)	(0.64, 0.76)	0.70 (0.19)	(0.61, 0.79)
Unfamiliar phone contrasts	0.58 (0.13)	(0.53, 0.63)	0.61 (0.16)	(0.53, 0.69)

Table 6. Results of fitting a taxonomy of mixed-effects logistic regression models for time 1 correct phonetic discrimination trials as a function of age, maternal education, phone type, language group and their interactions in a random sample of n = 60 preschoolers

Fixed	M1.1 β (se)	M1.2 β (se)	M1.3 β (se)
Intercept	.89 (.10)***	1.14 (.12)***	1.22 (.12)***
Age (in months, scaled)	.12 (.09)	.10 (.08)	.10 (.08)
Maternal education (0 = High School)			
1 = College	.12 (.19)		
2 = Graduate or Professional	.41 (.64)		
Phone type: (0 = familiar)			
1 = English-exclusive;	−.42 (.04) ***	−.46 (.04)***	−.47 (.06)***
2 = unfamiliar	−.65 (.04)***	−.67 (.04)***	−.82 (.05)***
Language: (0 = English DLL)			
1 = Spanish DLL		−.32 (.16)*	−.45 (.17)**
Interaction:			
Spanish DLL x Eng.-exclusive			.02 (.08)
Spanish DLL x unfamiliar			.27 (.07)***
Random			
N_{ID}	56	60	60
$Variance_{ID}$	0.39	0.39	0.39
Observations	1792	1920	1920
AIC	5907.8	6287.3	6273.1
deviance	5893.8	6277.3	6257.1

*p < .05, ** p < .01, *** p < .001
Note: Data on maternal education was missing for four children in the full Fall sample, leaving an effective sample size of 56 in M1.1.

significantly in their predicted probabilities of discriminating unfamiliar phones (post-hoc $\beta = -.18$, $p = .36$) at 60% for English- and 55% for Spanish-dominant DLLs.

Research question 2: language dominance and growth in phonetic discrimination

Next, we examined whether this language dominance by phone type interaction changed over time. When incorporating Spring testing, the baseline mixed-effects logistic regression model (table 7, Model 2.1) indicated that DLLs overall still had a significantly higher likelihood of correctly discriminating familiar compared to English-exclusive (odds = 1.5:1, $\beta = -.42$, $p < .001$) or unfamiliar (odds = 2.2:1, $\beta = -.81$, $p < .001$) phones. As seen in table 7, Model 2.2, DLLs improved in phonetic discrimination accuracy over time, with the odds of correct discrimination increasing by 1.57:1 ($\beta = .45$, $p < .001$) in Spring compared to Fall. Spanish- and English-dominant DLLs in Model 2.2 did not differ significantly in their likelihood

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

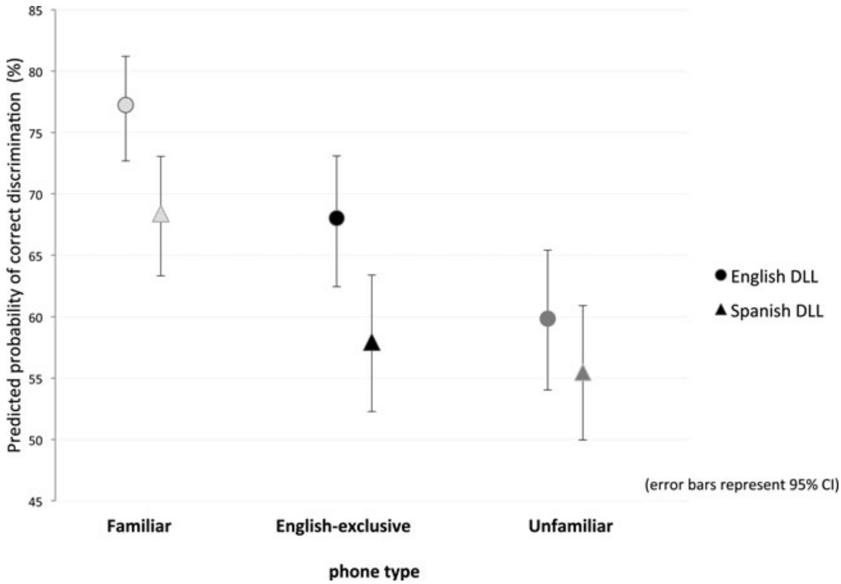

Figure 1. Model 1.3 results displaying predicted probabilities of time 1 correct phonetic discrimination by phone type & language group

of correct answers in the phone task overall (odds = 1.2:1, $\beta = -.20$, $p = .26$). However, there was a significant interaction between language group and phone type (table 7, Model 2.3): while English-dominant DLLs were still more likely than Spanish-dominant DLLs to correctly discriminate familiar phones (odds = 1.4:1, $\beta = -.37$, $p = .04$), groups no longer differed in discrimination of either English-exclusive (post-hoc $\beta = -.23$, $p = .26$) or unfamiliar (post-hoc $\beta = -.07$, $p = .79$) phones.

Consistent with these differences across Fall and Spring outcomes, Model 2.4 displays a significant three-way interaction between time, language group and phone type for English-exclusive phones ($\beta = .66$, $p < .001$). As illustrated in figure 2, while both groups grew across time in their ability to discriminate all phone types, by Spring, Spanish-dominant DLLs performed comparably to English-dominant DLLs in discriminating English-exclusive phones (post-hoc $\beta = -.02$, $p = 1.00$) with predicted probabilities of correct responses at 72% and 71% respectively. This model provided a significantly better fit to the data than the previous models when deviance and AIC were compared (all X^2, $p < .001$).

Research question 3: phonetic discrimination and preliteracy skills

We next considered whether phonetic discrimination in Fall predicted Spring phonological awareness and print knowledge, and whether these relationships were mediated by Spring phonetic discrimination performance. Since children's discrimination of unfamiliar phones was at or close to chance in Fall, we first considered the role of only familiar and English-exclusive phone discrimination in our baseline regression model predicting CELF-P2 Phonological Awareness performance. As in the prior research questions, we included age, maternal education

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

Table 7. Results of fitting a taxonomy of mixed-effects logistic regression models for correct phonetic discrimination trials as a function of age, maternal education, phone type, time, language group and their interactions, n = 42 preschoolers

Fixed	**M2.1** β (se)	**M2.2** β (se)	**M2.3** β (se)	**M2.4** β (se)
Intercept	1.05 (.11)***	1.02 (.14)***	1.14 (.15)***	.97 (.15)***
Age (in months) at time 2	.13 (.09)	.12 (.09)	.12 (.09)	.12 (.09)
Maternal education (0 = High School)				
1 = College	.02 (.21)			
2 = Graduate or Professional	.23 (.57)			
Phone type: (0 = familiar)				
1 = English-exclusive;	−.42 (.04)***	−.47 (.04)***	−.56 (.06)***	−.31 (.08)***
2 = unfamiliar)	−.81 (.03)***	−.86 (.03)***	−1.06 (.05)***	−.84 (.07)***
Time (Spring)		.45 (.03)***	.45 (.03)***	.86 (.08)***
Language: (0 = English DLL) 1 = Spanish DLL)		−.20 (.18)	−.37(.18)*	−.26 (.19)
Interaction:				
Spanish DLL x Eng.-exclusive			.13 (.07)	−.16 (.10)
Spanish DLL x unfamiliar			.29 (.06)***	.27 (.09)**
Interaction: Time x Spanish DLL				−.26 (.10)*
Interaction: Time x Eng.-exclusive				−.58 (.12)***
Time x unfamiliar				−.50 (.11)***
Interaction:				
Language x Time x Eng.-exclusive				.66 (.15)***
Language x Time x unfamiliar				.11 (.13)
Random				
N_{ID}	39	42	42	42
$Variance_{ID}$	0.31	0.29	0.29	0.29
Observations	2496	2688	2688	2688
AIC	8742.8	9118.0	9100.1	9033.3
deviance	8728.8	9104.0	9082.1	9005.3

*p < .05, **p < .01, ***p < .001

Note: Data on maternal education was missing for three children in the longitudinal sample, leaving an effective sample size of 38 in M2.1.

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

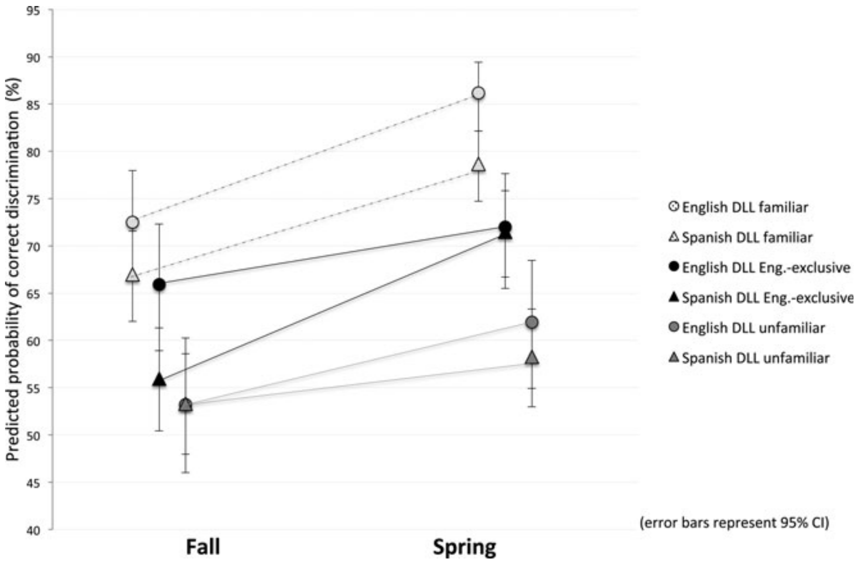

Figure 2. Model 2.4 results displaying predicted probabilities of correct phonetic discrimination by time, phone type & language group

and nonverbal reasoning as control variables in the baseline model but removed non-significant controls from subsequent models in order to conserve model degrees of freedom given the sample size. For this same reason, as the two phone-type coefficients were moderately positively correlated (Spearman's rho = 0.55, $p < .001$) in the baseline model, we did not subsequently disaggregate phonetic discrimination performance by phone type, and instead represented this performance by the task mean across all phone types (Fall phonetic discrimination). As seen in Table 8, Model 3.2, phonetic discrimination accuracy in Fall was a significant, positive predictor of CELF-P2 phonological awareness in Spring when controlling for age and maternal education ($\beta = 1.55$, $p = .02$). On average, there was no evidence of a difference between English- and Spanish-dominant DLLs on the English-language phonological awareness measure ($\beta = -.11$, $p = .59$). Similarly, when Spring discrimination accuracy was entered into the model (table 8, Model 3.3), the difference between English- and Spanish-dominant DLLs was not statistically significant ($\beta = -.04$, $p = .82$).

Spring and Fall phonetic discrimination were strongly correlated pairwise (Kendall's Tau=.366, $p = .001$). As seen in Model 3.3, the model coefficient for Fall discrimination was no longer statistically significant ($\beta = .44$, $p = .50$) when Spring discrimination was entered into the regression. Although our sample size did not allow for a full structural mediation model to reduce measurement error and account for directionality, this overlap of shared variance suggests that the relationship between Fall discrimination and Spring phonological awareness is mediated by Spring phonetic discrimination (Baron & Kenny, 1986).

Because Model 3.4 removing language group as a predictor was not significantly different from Model 3.3 (Likelihood ratio test $p = 0.78$), we selected Model 3.4 as the final model following the principle of parsimony. This final model demonstrated a

https://doi.org/10.1017/S0305000920000768 Published online by Cambridge University Press

Table 8. Results of fitting a taxonomy of generalized Poisson regression models for CELF-P2 Phonological Awareness raw scores as a function of age, maternal education, nonverbal reasoning, mean phonetic discrimination accuracy, and language group, n = 38 preschoolers

Fixed	M3.1 β (se)	M3.2 β (se)	M3.3 β (se)	M3.4 β (se)
Intercept	−3.07 (1.43)*	−2.40 (1.25)	−2.63 (1.07)*	−2.57 (1.07)*
Age (in months) in Spring	.05 (.02)*	.05 (.02)*	.04 (.02)*	.04 (.02)*
Maternal education (0 = High School)				
1 = College	.44 (.19)*	.43 (.23)	.39 (.19)*	.39 (.16)*
2 = Graduate or Professional (no observations)				
KBIT2 Nonverbal	.01 (.01)			
Time 1 phonetic discrimination (mean accuracy, Familiar phones)	.54 (.68)			
Time 1 phonetic discrimination (mean accuracy, Eng.-exclusive)	.58 (.68)			
Time 1 phonetic discrimination (mean accuracy, all trials)		1.55 (.64)*	.44 (.65)	
Time 2 phonetic discrimination (mean accuracy, all trials)			2.08 (.74)**	2.38 (.60)***
Language: (0 = English DLL) 1 = Spanish DLL		−.11 (.21)	−.04 (.18)	
N	38	38	38	38
Residual DOF	31	32	31	33
AIC	207.2	205.8	212.2	197.2
deviance	193.2	193.8	186.7	187.2

**p < .05, **p < .01, ***p < .001*

Note: Data on maternal education was missing for three out of 42 children in the longitudinal sample. One additional child did not complete the CELF-P2 Phonological Awareness subtest, leaving an effective sample size of 38 for all models.

significant, moderately-sized positive association between Spring phonetic discrimination and CELF-P2 phonological awareness. On average, a 10-percentage point difference in phonetic discrimination scores was associated with 27% additional items correct, or a relative incidence ratio of 1.27:1, on the CELF-P2 subtest ($\beta = 2.38$, $p < .001$) when holding age and maternal education constant.

As in the prior model taxonomy, our baseline model predicting print knowledge scores (Table 9, Model 4.1) included Fall discrimination of familiar and English-exclusive phones as the initial substantive predictors of interest with age, maternal education, and nonverbal reasoning as control variables. Maternal education ($\beta = 11.33$, $p = .11$) and nonverbal reasoning ($\beta = .07$, $p = .92$) were not statistically significant predictors of Spring print knowledge and were hence dropped from subsequent analysis. In contrast to the phonological awareness models, the two coefficients for our phonetic discrimination predictors for print knowledge differed in sign, with a negative coefficient for familiar phones and a positive coefficient for English-exclusive phones. Because this contrast suggested that the association of phonetic discrimination with print knowledge might be different for different phone types, we retained this variable disaggregated by type throughout the model taxonomy.

As seen in Table 9, Model 4.2, there was a significant positive association of Fall discrimination of English-exclusive ($\beta = 41.91$, $p = .046$) but not familiar ($\beta = -20.12$, $p = .29$) phones with print knowledge, when controlling for age and language group. Spanish-dominant DLLs performed on average lower than English-dominant DLLs across the model taxonomy (all $p < .05$) when controlling for age and phonetic discrimination performance. However, when Spring discrimination of English-exclusive phones ($\beta = 54.34$, $p = .001$) was entered into the regression (table 9, Model 4.3), the Fall English-exclusive measure was no longer statistically significant ($\beta = 25.40$, $p = .09$). Just as with phonological awareness (table 8), Spring phonetic discrimination appears to mediate the relationship between Fall discrimination and Spring print knowledge. Model 4.4, the final best-fitting and most parsimonious model, included the main effects of child age, English-exclusive phone discrimination in Spring, and language group. When holding age and phonetic discrimination constant, Spanish-dominant DLLs were predicted to score on average 14.4 points lower ($\beta = -14.30$, $p = .003$) than English-dominant DLLs on the Print Knowledge subtest, a difference of almost one standard deviation. English-exclusive phone discrimination displayed a strong but relatively modest association with Print Knowledge performance. On average, each 10-percentage point positive difference in English-exclusive phone discrimination was associated with a 5-point, or one-third standard deviation, positive difference in Print Knowledge subtest scores ($\beta = 56.18$, $p < .001$) when controlling for age and language group.

Discussion

Our findings contribute to the growing body of literature that indicates that bilinguals are a highly diverse group (Surrain & Luk, 2017) and this heterogeneity impacts language outcomes among Spanish–English DLLs (Kim, Richards & Burts, 2018). In the current study, DLL within-population diversity related to language dominance showed a relationship with differential development of English phonetic discrimination. Additionally, phonetic discrimination accuracy was a significant, positive predictor of preliteracy skills (phonological awareness and print awareness) in the Spring.

Table 9. Results of fitting a taxonomy of truncated Gaussian regression models for TOPEL Print Knowledge as a function of age, nonverbal reasoning, mean phonetic discrimination accuracy, and language group, n = 41 preschoolers

Fixed	M4.1 β (se)	M4.2 β (se)	M4.3 β (se)	M4.4 β (se)
Intercept	157.65*** (41.32)	152.73*** (34.17)	125.06*** (24.72)	118.03*** (26.26)
Age (in months) in Spring	−1.27* (.61)	−1.11 (.55)	−1.09* (.41)	−1.03* (.43)
Maternal education (0 = High School)				
1 = Undergraduate	11.33 (6.9)			
2 = Graduate or Professional (no observations)				
KBIT2 Nonverbal	.02 (.22)			
Time 1 phonetic discrimination (mean accuracy, Familiar phones)	−22.85 (21.88)	−20.12 (18.80)	−28.51 (15.10)	
Time 1 phonetic discrimination (mean accuracy, Eng.-exclusive)	52.67* (22.84)	41.91* (20.26)	25.40 (14.54)	
Time 2 phonetic discrimination (mean accuracy, Familiar phones)			3.09 (13.00)	
Time 2 phonetic discrimination (mean accuracy, Eng.-exclusive)			54.34** (15.53)	56.18*** (14.40)
Language:				
(0 = English DLL; 1 = Spanish DLL)		−14.2* (5.92)	−12.62** (4.36)	−14.30** (4.56)
N	38	41	41	41
Residual DOF	31	35	33	36
AIC	300.8	318.0	305.9	304.5
deviance	286.8	306.0	289.9	294.5

* *p* < .05, ** *p* < .01, *** *p* < .001

Note: Data on maternal education was missing for three out of 42 children in the longitudinal sample. One additional child did not complete the TOPEL Print Knowledge subtest, leaving an effective sample size of 38 for the Time 1 baseline model (M4.1) and 41 for the subsequent longitudinal models.

Three main findings were observed in the present study. First, we found that Spanish-dominant and English-dominant DLLs had differential discrimination for English phones that are familiar and English-exclusive, but not for unfamiliar phones during the first time of assessment. Second, this differential discrimination changes in a 4–6 month span such that language dominance did not modulate discrimination accuracy in English-exclusive phones. Third, phonetic discrimination, but not language dominance, predicted phonological awareness, while both discrimination and dominance predicted print knowledge in English. Further, phonetic discrimination in Spring may mediate the relationship between Fall discrimination and both of these pre-literacy skills. Each of these findings are considered in turn.

The results from our study suggest that phonemic representation may begin to emerge driven by lexical restructuring. It is, however, possible that representations might be further sharpened through literacy instruction (Werker & Curtin, 2005). Our findings reinforce two converging messages for research and practice among preschool DLLs: it is necessary to recognize the heterogeneity within Spanish–English bilingual children categorized by the U.S. education system as DLLs; and base judgements of preschool-aged DLLs' learning outcomes on repeated assessments, particularly those recently exposed to an intensive English environment.

Emergent DLL preschoolers showed differential phonetic discrimination; as hypothesized, all participants better discriminated between English phones with Spanish analogues than English-exclusive phones. Children identified as English-dominant DLLs likely had more English exposure and performed more accurately than Spanish-dominant DLL peers on English-exclusive phones in Fall. Spanish-dominant DLLs performed close to chance (BCa bootstrap interval 0.52–0.63) when discriminating between English-exclusive phones at Fall testing but were comparable to English-dominant DLLs when tested again in Spring, 4–6 months later, indicating rapid changes in discrimination of English-exclusive phones. We conjecture that Spanish-dominant DLLs would outperform English-dominant DLLs if given a comparable measure of Spanish phones with and without English analogues. Both types of English phones were better recognized at time 2 when all participants had more exposure to English. As expected, all participants were not as able to discriminate unfamiliar phones (e.g., Russian phones not present in English or Spanish) and performance on these items did not differ significantly from chance in initial Fall testing and displayed the smallest growth across the two time points compared to the other phones.

These findings align with the broader understanding that phonological boundaries are influenced by language experiences after the initial period of perceptual narrowing. Our study indicates that DLL preschoolers likely continue to refine representations for more detail, including when exposed to new language environments with different phonemes. Our findings demonstrate that children's discrimination abilities increase over time, although it should be noted that the current study does not use a direct measure of SEGMENTAL representation. We interpret accurate discrimination of the pseudo-words used in the task as evidence of forming phonemic representations; however, currently the task does not reveal if fully phonemic representations have developed.

Bialystok and colleagues (2010) described the dual language development of bilingual children as "constructing the world through two telescopes" and suggested that "their two vocabularies provide the lenses" (p.530). Building on this metaphor, we posit that children construct phonological representations by "viewing" the sounds of speech through these two lenses. As bilingual children construct the world through two lexical telescopes, they simultaneously refine and reorganize existing representations. As their expanding lexicon includes progressively more English words, DLL children likely access existing representations for further refinement, such as updating for phonemes exclusive to English.

Performance on our phonetic discrimination task also showed significant relationships with other measures associated with phonological sensitivity and/or language. Specifically, as demonstrated in the model taxonomy (Models 3.1–3.4) for phonological awareness, overall phonetic discrimination performance in Fall predicted more conventional phonological awareness assessments administered in Spring. Further, these relationships between Fall phonetic discrimination and Spring pre-literacy tasks appears to be mediated by Spring phonetic discrimination. This mediation suggests that the phonetic discrimination task overall, while not requiring specific vocabulary or letter-sound knowledge, nevertheless taps into concurrent phonological abilities. Mediation may also indicate the stability of the importance of phonetic discrimination in phonological awareness development at this age.

Additionally, English-exclusive phone performance in Fall showed a significant relationship with performance on the TOPEL Print Knowledge (an English assessment) in Spring. This relationship was not observed between familiar phones discrimination and print knowledge. This finding suggests the role of English-exclusive phones in relation to English literacy; it may be that the ability to discriminate between English-exclusive sounds in particular supports the language skills directly or indirectly captured by assessments of English print knowledge, like the TOPEL, or that English-exclusive phones serve as a proxy (i.e., children who have had more explicit literacy instruction may have higher performance on both measures). We suggest that future research account for variables such as amount of explicit literacy instruction. Future research should examine the importance of discriminating between English-exclusive phones in particular and explore its predictive power of subsequent literacy skills.

The developing phonological system

The current study also provides insight into how phonological representations develop for DLLs. Our findings indicate that DLLs may initially struggle to discriminate between phones that they have not been intensively exposed to in the home, but after 4–6 months' additional preschool English exposure, they are able to more accurately make judgments. These findings further support our understanding that the phonological system changes even at age 4 to 5 years, after the hypothesized period of perceptual narrowing, for children developing in multiple languages. We interpret our findings as further evidence of a distinct developmental pattern for DLLs and extended timelines for establishing native categories (see Birdsong, 2018). Observed performance changes between Fall and Spring testing on English-exclusive and familiar items among Spanish-dominant DLLs is potential evidence of changes in phonological representation boundaries.

Our findings lend support for theories of phonological development that propose that phonemic representation begins to emerge driven by lexical restructuring (e.g., Lexical restructuring model plus letters; Ventura et al., 2007). Phonemic representations could be further sharpened through literacy instruction in line with alternate models (Werker & Curtin, 2005). Our measures associated with English literacy instruction (English phonological awareness, English print knowledge) were only administered in Spring after all DLLs had at least 4–6 months of English instruction. These tasks rely on English word knowledge and have not been created for DLLs without or with minimal English-language experience: we therefore considered the measures inappropriate for our participants in Fall. As such, we are only able to examine relationships in a single direction (phonetic discrimination to phonological awareness and print knowledge), and we are unable to explore potential bidirectionality of the relationships between these variables. Additionally, prior research has suggested that phonemic representations may be only fully realized when children develop explicit awareness through literacy instruction (Carroll, 2004; for contrasting perspectives see Ainsworth et al., 2019; Ventura et al., 2007). It is therefore possible that participants' improved accuracy discriminating between English exclusive and familiar phones was related to classroom literacy instruction. The current study cannot address questions related to whether or not children possess detailed phonological representations that are accessed only once they have metacognitive and letter knowledge (i.e., Accessibility Account; Liberman et al., 1989). This is a potential direction for future research. Our findings of differential phonetic discrimination of English-exclusive phones and English phones with Spanish analogues can also inform understanding of the unique development of the phonological system and phonological awareness among DLLs in the U.S., in particular as it relates to transfer between Spanish and English. The body of research on cross-linguistic phonological knowledge remains contradictory and many factors relevant to relationships between phonological skills in two languages have not been fully explored (e.g., Gottardo, Gu, Mueller, Baciu & Pauchulo, 2011). Evidence generally supports a bidirectional relationship between L1 and L2 phonological awareness among Spanish–English DLL preschoolers. DLLs with high phonological awareness in Spanish also tend to have high English phonological awareness (for a review, see Soto et al., 2019). Furthermore, DLLs who receive phonological awareness instruction in both Spanish and English improve phonological awareness in both languages and make greater gains in English phonological awareness than DLLs who receive instruction in only English (Soto et al., 2019). The current study examined discrimination of English-exclusive phones (i.e., those with minimal potential for transfer) and English phones with Spanish analogues (i.e., those with potential for transfer) among DLLs. Our findings indicate that there may be differential acquired ability to discriminate between English phones with and without Spanish analogues. Preschool DLLs may be refining their "lenses" and, as such, phonological awareness instruction could include both languages to support developing representations of English phones with Spanish analogues, via potential positive transfer.

Receptive measurement of phonetic discrimination

The current study presents a novel measure of phonetic discrimination with English phones that do and do not have analogues in Spanish. To our knowledge, it is the first receptive measure for preschool-aged Spanish–English bilinguals that explicitly separates and compares English phones with and without Spanish analogues (along with unfamiliar controls) thus facilitating examination of differential changes in discrimination of these phones during the preschool years. The phonetic discrimination task takes less than 10 minutes to complete and is easily administered via touch-screen tablet. Additionally, it does not require tester knowledge of Spanish.

The simple instructions are presented with visual supports and pre-task training ensures understanding of match/mismatch. We think this receptive measure of English phonetic discrimination has the potential to contribute to the growing body of research on Spanish–English bilingual children by providing broad insights into phonological development among bilingual learners.

There is ambiguity regarding whether our phonetic discrimination task addresses phones or phonemes, specifically if accurate discrimination between pseudo-words that begin with English-exclusive and familiar phones indicates phonemic representations. We interpret discriminating pseudo-words as evidence of an emergent language-specific phoneme representation. However, we acknowledge that our task does not reveal if fully phonemic representations have developed.

DLL heterogeneity

Previous studies have found that DLL within-population diversity impacts language outcomes. Kim and colleagues (2018) used latent profile analysis to identify three distinct DLL subgroups among preschoolers: emergent bilinguals (children who spoke L1 and some English in the home and classroom); bilinguals (children who spoke L1 and English in the home, only English in the classroom), and heritage language speakers (children who spoke only the L1 in the home and classroom). These three distinct demographic profiles were associated with different development and learning trajectories across various outcomes (language, literacy, and mathematics) over a period of one school year. In particular, the "bilingual" DLL subgroup showed fewer performance gaps when compared to non-DLL peers, and discrepancies that did exist became smaller over time. DLL subgroups also differed from each other, as well as from non-DLL peers, on cognitive, social-emotional, and motor/physical outcomes. The "bilingual" subgroup outperformed non-DLL peers in all three of the above areas over the duration of the study. In the beginning of the year, "emergent bilingual" and "heritage language" subgroups performed more weakly than non-DLLs, yet ultimately surpassed non-DLL peers by the end (Kim, Richards & Burts, 2018). These findings of relevant differences and differences over time, with regards to cognitive outcomes, are in line with previous research indicating that certain forms of bilingual engagement during childhood are associated with cognitive advantages (Bialystok, 2018; Hartanto, Toh & Yang, 2018; Santillán & Khurana, 2017). Similarly, Foster and Anthony (2019) used latent profile analysis to examine mathematics achievement among over 500 DLL kindergarteners and identified four distinct profiles ("Spanish-dominant", "balanced high language proficiency", "English-dominant", "balanced low language proficiency"), each associated with different mathematics development trajectories.

Our findings regarding subgroup differences in knowledge of English phones expands the existing body of research on DLL subgroup development and makes a specific contribution addressing discrimination of English phones. Given the importance of phonological awareness for English reading outcomes, our findings of subgroup differences have implications for practitioners. The distinction between DLLs who are Spanish-dominant versus those who are English-dominant, as defined in the current study, thus has implications for expected growth profiles and changes over time.

Limitations and directions for future research

The current study does not present a complete picture of Spanish-dominant or English-dominant DLL preschoolers' phonological systems. The phonetic discrimination measure used does not address Spanish phones or discrimination of Spanish phones. The measure is only intended to measure discrimination between phones relevant for English and, relatedly, English literacy. In English, knowledge of and ability to manipulate phones at the sound level is essential for reading success, while in Spanish, blending and segmenting syllables better predicts Spanish literacy achievement (Soto et al., 2019). As such, our phonetic discrimination task designed to specifically measure English phonetic discrimination among Spanish–English DLLs does not inform on the development of Spanish, which is relevant and necessary for understanding the overarching, unique development of phonological awareness among Spanish–English bilingual children. A Spanish phone discrimination task that emphasizes contrasts exclusive to Spanish (for example, trilled r) and familiar Spanish phones with an English analogue could be developed. A receptive measure of Spanish blending and segmenting could provide even further insight into the development of DLLs' phonological system and potential shifts in Spanish phone discrimination after English exposure.

Additionally, there are many fundamental aspects of phonological awareness and phonological system development beyond onset consonant-level discrimination and, as such, implications from the current study are limited. Soto and colleagues (2019) note that emphasis on discriminating between small speech segments (phonemes) aligns with a phonological system development trajectory from larger to smaller segments of speech (i.e., syllable- and phoneme-level knowledge), which is generally held to be the trajectory for English. The developmental trajectory for Spanish phonological awareness, however, progresses from the ability to manipulate two-syllable words to increasingly long ones (Soto et al., 2019). It is therefore also possible that the Spanish–English DLLs were simply less familiar with this type of phonological discrimination (CVC onset comparison), although they did perform above chance when discriminating familiar phones, indicating at least some ability.

Another limitation of the phonetic discrimination task used in the current study is the need for task instructions in the child's preferred language. It is possible that by presenting task instructions in Spanish for the children identified as Spanish–English DLLs, learners were primed for Spanish and then had to "shift" when attending to the English speech sounds that followed (CVC pseudo-words, phones unique to English). Given that it was essential that the child understand the task instructions and demonstrate the ability to perform the task, this was considered an acceptable limitation. No further narration in Spanish or English was given once the child demonstrated the ability to successfully complete training trails.

The sample in the current study has limitations and as such findings should not be generalized to all Spanish–English DLLs or other bilinguals more broadly. One key limitation of the current study sample is the significant difference in maternal education background and a trend toward difference in household income between the two language dominance groups. Socio-economic status (SES) has generally been shown to correlate strongly with general language development and higher-level language skills (i.e., vocabulary) in childhood (Pace, Luo, Hirsh-Pasek & Golinkoff, 2017). Thus, it is possible that differences in measure performance between the two language-dominance groups are also related to SES. Additionally, research with bilingual populations has found that bilingual language environments may result in refinements in phonological processing in older children and adults even in lower SES contexts (Krizman, Skoe & Kraus, 2016; Skoe, Burakiewicz, Figueiredo & Hardin, 2017). However, the relationship among complex and simultaneously occurring environmental factors such as socioeconomic status and bilingualism is still unclear. We accept this limitation as reflective of the population where the sample was collected. A larger sample would allow for further controlling of background variables related to SES, as well as permitting the use of latent profile analysis to describe more than two DLL subgroup profiles relevant to English phone discrimination. Finally, regression modeling has intrinsic measurement limitations. A larger sample would allow for the development of a measurement model using structural equation modeling, creating a more precise and reliable estimation of coefficients, providing evidence of directionality and mediation, and better separating the variance of predictive measures.

The current study underscores the perils of approaching "DLL" or bilingual as a categorical variable (Luk & Bialystok, 2013) and the importance of longitudinal research when examining preschool DLL development (Hammer, Jia & Uchikoshi, 2011; Kim, Lampert & Burt, 2018). Practitioners can better meet DLLs' needs if they have an understanding of DLL within-population diversity and the potential for the relevance of distinct subgroup profiles. As such, future research should continue to identify and share information regarding DLL heterogeneity, including describing subgroup profiles relevant to various outcome variables (i.e., for mathematics achievement, there may exist four subgroups; for cognitive outcomes, three; for discrimination of English phones, perhaps two). Our findings could have implications for school professionals. A next step may be to examine how school professionals view DLLs, i.e., as a heterogeneous or homogeneous group, and, depending on findings, inform via explicit training both pre- and in-service teachers. Optimal phonological awareness instruction likely requires differentiation and future research could focus on methods and measures to help teachers to reliably identify DLL-relevant subgroups and make aligned instructional choices.

Acknowledgements. We would like to thank all the children, families, teachers, and preschools who participated in this study. We also thank the anonymous reviewers and the Action Editor for their comments that greatly improved the manuscript. This project is partially funded by the Bezos Family Foundation and the National Academy of Education/Spencer Foundation postdoctoral fellowship to the last author.

References

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print*. Cambridge, MA: MIT Press.

Ainsworth, S., Welbourne, S., & Hesketh, A. (2016). Lexical restructuring in preliterate children: Evidence from novel measures of phonological representation. *Applied Psycholinguistics*, *37*(4), 997–1023.

Ainsworth, S., Welbourne, S., Woollams, A., & Hesketh, A. (2019). Contrasting explicit with implicit measures of children's representations: The case of segmental phonology. *Language Learning*, *69*(2), 323–65.

Ambrose, S. E., Fey, M. E., & Eisenberg, L. S. (2012). Phonological awareness and print knowledge of preschool children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *55*(3), 811–23.

Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, *46*(2–3), 155–82.

Anthony, J. L., Solari, E. J., Williams, J. M., Schoger, K. D., Zhang, Z., Branum-Martin, L., & Francis, D. J. (2009). Development of bilingual phonological awareness in Spanish-speaking English language learners: The roles of vocabulary, letter knowledge, and prior phonological awareness. *Scientific Studies of Reading*, *13*(6), 535–564.

Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 699–714.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bialystok, E. (2018). Bilingualism and executive function. *Bilingual Cognition and Language: The state of the science across its subfields*, *54*, 283–306.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–31.

Birdsong, D. (2018). Plasticity, variability and age in second language acquisition and bilingualism. *Frontiers in psychology*, *9*, 81, doi: 10.3389/fpsyg.2018.00081.

Bosch, L., & Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and speech*, *46*(2–3), 217–43.

Branum-Martin, L., Tao, S., Garnaat, S., Bunta, F., & Francis, D. J. (2012). Meta-analysis of bilingual phonological awareness: Language, age, and psycholinguistic grain size. *Journal of Educational Psychology*, *104*(4), 932.

Burns, T. C., Yoshida, K. A., Hill, K., & Werker, J. F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, *28*(3), 455–74.

Bus, A. G., & Van IJzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of educational psychology*, *91*(3), 403–14.

Bus, A. G., Van Ijzendoorn, M. H., & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. Review *of Educational Research*, *65*(1), 1–21.

Byers-Heinlein, K., & Fennell, C. T. (2014). Perceptual narrowing in the context of increased variation: Insights from bilingual infants. *Developmental Psychobiology*, *56*(2), 274–91.

Byrne, B. J. (2014). *The foundation of literacy: The child's acquisition of the alphabetic principle* (second edition). Psychology Press.

Carroll, J. M. (2004). Letter knowledge precipitates phoneme segmentation, but not phoneme invariance. *Journal of Research in Reading*, *27*(3), 212–225.

Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. J*ournal of Child Psychology and Psychiatry*, *45*(3), 631–40.

Curenton, J. M., & Justice, L. M. (2008). Children's preliteracy skills: Influence of mothers' education and beliefs about shared reading interactions. *Early Education and Development*, *19*, 261–283.

Curtin, S., & Zamuner, T. S. (2014). Understanding the developing sound system: Interactions between sounds and words. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(5), 589–602.

Eilers, R. E., Wilson, W. R., & Moore, J. M. (1977). Developmental changes in speech discrimination in infants. *Journal of Speech and Hearing Research*, *20*(4), 766–780.

Enriquez, B. E. (2015). *Head Start Early Learning Outcomes Framework: Ages Birth to Five*. U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/elof-ohs-framework.pdf

Fennell, C. T., Tsui, A. S. M., & Hudon, T. M. (2016). Speech perception in simultaneously bilingual infants. *Bilingualism across the lifespan: Factors moderating language proficiency*. Washington, USA: American Psychological Association.

Ferjan Ramírez, N., Ramírez, R. R., Clarke, M., Taulu, S., & Kuhl, P. K. (2017). Speech discrimination in 11-month-old bilingual and monolingual infants: a magnetoencephalography study. *Developmental Science*, *20*(1), 1–16.

Foster, M. E., & Anthony, J. L. (2019). English and Spanish profiles of oral language growth trajectories of dual language learners and monolingual English speakers. Paper presented at the annual meeting of the Pacific Coast Research Conference. San Diego, CA.

Fowlert, A. E., Swainson, B., & Scarborough, H. (2004). Relationships of naming skills to reading, memory, and receptive vocabulary: Evidence for imprecise phonological representations of words by poor readers. *Annals of Dyslexia*, *54*(2), 247.

Garcia-Sierra, A., Ramírez-Esparza, N., & Kuhl, P. K. (2016). Relationships between quantity of language input and brain responses in bilingual and monolingual infants. *International Journal of Psychophysiology*, *110*, 1–17.

Genesee, F. (1989). Early bilingual development: One language or two?. *Journal of child language*, *16*(1), 161–179.

Gildersleeve, C., Davis, B., & Stubbe, E. (1996, November). When monolingual rules don't apply: Speech development in a bilingual environment. In American Speech-Language-Hearing Association Convention, Seattle, WA.

Gildersleeve-Neumann, C., & Davis, B. (1998). Learning English in a bilingual preschool environment: Change over time. In *Annual Convention of the American Speech-Language-Hearing Association*, San Antonio, TX.

Goldstein, B., & Washington, P. S. (2001). An initial investigation of phonological patterns in typically developing 4-year-old Spanish–English bilingual children. *Language, Speech, and Hearing Services in Schools*, *32*(3), 153–164.

Gonzales, K., Byers-Heinlein, K., & Lotto, A. J. (2019). How bilinguals perceive speech depends on which language they think they're hearing. *Cognition*, *182*, 318–330.

Gottardo, A., Gu, Y., Mueller, J., Baciu, I., & Pauchulo, A.L. (2011). Factors affecting the relative relationships between first- and second-language phonological awareness and second language reading. In A. Y. Durgunoglu (Ed.), *Language and literacy development in bilingual settings* (pp. 141–167). New York: Guilford.

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 715–733.

Hammer, C. S., Jia, G., & Uchikoshi, Y. (2011). Language and literacy development of dual language learners growing up in the United States: A call for research. *Child Development Perspectives*, *5*(1), 4–9.

Hammer, C. S., & Miccio, A. W. (2006). Early language and reading development of bilingual preschoolers from low-income families. *Topics in Language Disorders*, *26*(4), 322.

Hartanto, A., Toh, W. X., & Yang, H. (2018). Bilingualism Narrows Socioeconomic Disparities in Executive Functions and Self-Regulatory Behaviors During Early Childhood: Evidence From the Early Childhood Longitudinal Study. *Child development*, *90*(4), 1215–35.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research*, *23*(1), 65–94.

Hoff, E. (2018). Bilingual development in children of immigrant families. *Child development perspectives*, *12*(2), 80–86.

Jiménez González, J. E., & García, C. R. H. (1995). Effects of word linguistic properties on phonological awareness in Spanish children. *Journal of Educational Psychology*, *87*(2), 193.

Kaufman, A. S. (2004). *KBIT-2: Kaufman Brief Intelligence Test. 2nd*. Circle Pines, Minn: American Guidance Service.

Kim, D. H., Lambert, R. G., & Burts, D. C. (2018). Are young dual language learners homogeneous? Identifying subgroups using latent class analysis. *The Journal of Educational Research*, *111*(1), 43–57.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental science*, *9*(2), F13-F21.

Krizman, J., Skoe, E., & Kraus, N. (2016). Bilingual enhancements have no socioeconomic boundaries. *Developmental Science*, *19*(6), 881–891.

Leon Guerrero, S., Smith, S., & Luk, G. (2016). Home language usage and executive function in bilingual preschoolers. *Cognitive Control and Consequences of Multilingualism. John Benjamins Publishing Company*, *2*, 351.

Liberman, I. Y., Shankweiler, D., & Liberman, A. M. (1989). The alphabetic principle and learning to read. In *Meeting of the International Academy for Research on Learning Disabilities*, Oct, 1986, Northwestern U, Evanston, IL, US. The University of Michigan Press.

Lonigan, C. J., Schatschneider, C., Westberg, L., & others. (2008). Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. *Developing Early Literacy: Report of the National Early Literacy Panel*, 55–106.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, *25*(5), 605–621.

Mattys, S. L., Jusczyk, P. W., Luce, P. A., & Morgan, J. L. (1999). Phonotactic and prosodic effects on word segmentation in infants. *Cognitive psychology*, *38*(4), 465–494.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restrucuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Hillsdale, NJ: Erlbaum.

Milburn, T. F., Lonigan, C. J., & Phillips, B. M. (2019). Stability of Risk Status During Preschool. *Journal of Learning Disabilities*, *52*(3), 209–219. https://doi.org/10.1177/0022219418789373

Office of Head Start (2018). *Office of Head Start services snapshot: National all programs (2017–2018).* Office of Head Start: Washington, DC, USA.

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, *3*, 285–308.

Paradis, J., Nicoladis, E., Crago, M., & Genesee, F. (2011). Bilingual children's acquisition of the past tense: A usage-based approach. *Journal of Child Language*, *38*(3), 554–578.

Park, M., Zong, J., & Batalova, J. (2018). *Growing superdiversity among young US dual language learners and its implications*. Washington, DC: Migration Policy Institute.

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in neuroinformatics*, *2*, 10.

Pierce, L. J., Genesee, F., Delcenserie, A., & Morgan, G. (2017). Variations in phonological working memory: Linking early language experiences and language learning outcomes. *Applied Psycholinguistics*, *38*(6), 1265–1300.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of/d/–/ð/perception: evidence for a new developmental pattern. *The Journal of the Acoustical Society of America*, *109*(5), 2190–2201.

Quay, S., & Montanari, S. (2016). Early bilingualism: From differentiation to the impact of family language practices. *Bilingualism across the lifespan*, *23–42*.

R Core Team. (2018). *R: A language and environment for statistical computing; 2015*.

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape, (with discussion). *Applied Statistics*, *54*, 507–554.

Rowe, M. L., Denmark, N., Harden, B. J., & Stapleton, L. M. (2016). The role of parent education and parenting knowledge in children's language and literacy skills among White, Black, and Latino families. *Infant and Child Development*, *25*(2), 198–220.

Rvachew, S., Ohberg, A., Grawburg, M., & Heyding, J. (2003). Phonological awareness and phonemic perception in 4-year-old children with delayed expressive phonology skills. *American Journal of Speech-Language Pathology*, *12*(4), 463–471.

Sadeghi, A., & Everatt, J. (2017). Phonological Awareness Development in Speakers of Languages Other Than English. In G. Gillon (Ed.) *Phonological Awareness, Second Edition: From Research to Practice*. New York: Guilford.

Santillán, J., & Khurana, A. (2017). Developmental associations between bilingual experience and inhibitory control trajectories in Head Start children. *Developmental science*.

Schnitzer, M.L. (1997). *Spanish–English Contrastive Phonology*. San Juan, PR: Piedras Press, Inc.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*(2), B11–B21

Skoe, E., Burakiewicz, E., Figueiredo, M., & Hardin, M. (2017). Basic neural processing of sound in adults is influenced by bilingual experience. *Neuroscience*, *349*, 278–290.

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, *7*(1), 37–46.

Soto, X., Olszweski, A., & Goldstein, H. (2019). Systematic review of phonological awareness interventions for Latino children in early and primary grades. *Journal of Early Intervention*.

Storch, S.A., & Whitehurst, G.J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, *38*(6), 934.

Surrain, S., & Luk, G. (2017). Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015. *Bilingualism: Language and Cognition*, 1–15.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*(2), 147–166.

Torgesen, J.K. (2004). Avoiding the devastating downward spiral: The evidence that early intervention prevents reading failure. *American Educator*, *28*(3), 6–19.

Ventura, P., Kolinsky, R., Fernandes, S., Querido, L., & Morais, J. (2007). Lexical restructuring in the absence of literacy. *Cognition*, *105*, 334–361.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language learning and development*, *1*(2), 197–234.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, *103*(1), 147–162.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and development*, *7*(1), 49–63.

Wiig, E. H., Secord, W., & Semel, E. M. (2004). *CELF preschool 2: clinical evaluation of language fundamentals preschool.* Pearson/PsychCorp.

Unsworth, S. (2016). Early child L2 acquisition: Age or input effects? Neither, or both?. *Journal of child language*, *43*(3), 608–634.

Cite this article: Smith SA, Leon Guerrero S, Surrain S, Luk G (2022). Phonetic discrimination, phonological awareness, and pre-literacy skills in Spanish–English dual language preschoolers. *Journal of Child Language* **49**, 80–113. https://doi.org/10.1017/S0305000920000768