Research Article

Semantic Category Convergence in Spanish–English Bilingual Children With and Without Developmental Language Disorder

Prarthana Shivabasappa,^a Elizabeth D. Peña,^b and Lisa M. Bedore^c

Purpose: The study examines the extent of convergence of semantic category members in Spanish–English bilingual children with reference to adults using a semantic fluency task.

Method: Thirty-seven children with developmental language disorder (DLD), matched pairwise with 37 typically developing (TD) children in the age range of 7;0–9;11 (years;months), produced items in 7 semantic categories (3 taxonomic and 4 slot-filler) in both Spanish and English. The 10 most frequently produced items for each category by 20 Spanish–English bilingual adults were identified as the most prototypical responses. The top 10 items generated by TD children and children with DLD, in their order of production, were analyzed for the amount of convergence with adults' responses.

Results: The top 5 items produced by children with DLD showed similar convergence scores as those produced

by their TD peers. However, their responses in the 6th to 10th positions showed lower convergence scores than their TD peers. Children's convergence scores were higher for the slot-filler condition compared to taxonomic in both English and Spanish. The convergence scores also significantly differed across the semantic categories.

Conclusion: The children with DLD show greater convergence on the typical items generated earlier in their word lists than the items generated later. This pattern of convergence and divergence highlights their strengths and weaknesses in the representation of lexical–semantic knowledge for typical versus less typical items.

Supplemental Material: <https://doi.org/10.23641/asha.8323613>

This study aimed to investigate semantic category structure in bilingual children with developmental language disorder (DLD) and typically developing (TD) children and to compare their semantic convergence relative to bilingual adults. *Semantic convergence* refers to the extent to which patterns of word use approximate the patterns used by adults with similar cultural and linguistic backgrounds (Adams & Bullock, 1986). Comparing children's convergence with adult patterns of use informs our understanding of children's mastery of normative conditions of use by taking into account the conventional cultural realizations of concepts. The process of semantic development relies heavily on learning conventional use of linguistic

patterns (Clark, 2007). Semantic convergence is also crucial for development of in-depth word knowledge facilitating robust semantic networks. Thus, focusing on convergence allows us to evaluate individual sensitivity to words available in the environment.

Semantic category fluency tasks are useful in examining how children and adults access their lexicon, represent, and organize semantic categories (Hedden, Lautenschlager, & Park, 2005; Hughes & Bryan, 2002; Hurks et al., 2010; Riva, Nichelli, & Devoti, 2000). In this task, the participant is typically asked to list items belonging to a category (e.g., animal) for 60 s. Participants must strategically explore their lexical–semantic network to access relevant exemplars that align with the category while inhibiting the exemplars that are outside the category or already produced.

TD children acquire adultlike categorization skills with age resulting from increased language experience (Bjorklund, Thompson, & Ornstein, 1983). The lexical–semantic deficits of children with DLD are manifested during category fluency tasks as difficulties in recalling of appropriate category members (Coelho, Albuquerque, & Simões, 2013). They often produce fewer items, make a greater number of errors, and have longer latencies compared to TD peers (Hall,

^aDepartment of Communication Sciences & Disorders, The University of Texas at Austin

^bSchool of Education, University of California, Irvine

^cDepartment of Communication Sciences and Disorders, Temple University, Philadelphia, PA

Correspondence to Prarthana Shivabasappa: prarthana@utexas.edu

Editor-in-Chief: Sean Redmond

Editor: Lizbeth Finestack

Received November 14, 2017

Revision received April 25, 2018

Accepted January 27, 2019

https://doi.org/10.1044/2019_JSLHR-L-17-0427

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

McGregor, & Oleson, 2017; Weckerly, Wulfeck, & Reilly, 2001). In a bilingual scenario, children reaching adultlike skills are dependent on the amount and nature of exposure they receive in each language. Bilingual children with DLD demonstrate lexical–semantic deficits such as limited vocabulary and slow and inefficient retrieval of previously acquired words (Kohnert & Kan, 2007; Restrepo & Kruth, 2000). They also present with reduced semantic depth compared to their TD matched peers reflecting sparse semantic links and less robust semantic representation (Sheng, Peña, Bedore, & Fiestas, 2012). Limited semantic knowledge negatively impacts children’s functional communication skills and academic success (Uccelli, Galloway, Barr, Meneses, & Dobbs, 2015), whether the source of variation is bilingualism or language disorder.

One way to determine the extent of semantic learning and organization is by examining convergence patterns of children’s productions in category fluency tasks. By exploring how bilingual children with DLD compare with their TD peers in category fluency tasks, we will gain insight into how children with DLD retrieve and organize semantic knowledge. Semantic convergence has, to our knowledge, been studied only sparsely in children with DLD. One study focuses on convergence patterns of bilingual children with and without DLD. Sheng, Bedore, Peña, and Taliancich-Klinger (2013) examined semantic convergence of repeated word association responses in Spanish–English bilingual children with and without DLD. TD children produced responses that were similar in frequency to those produced by a normative group of TD children. Children with DLD, on the other hand, demonstrated less semantic convergence, although, as a group, they produced many of the same items, the rank frequencies were below those of TD children. An open question then is at what point in a semantic fluency task do children with DLD diverge from their TD peers.

Category Fluency Task

Children initially categorize items that occur in their immediate environment using slot-filler or script-based strategies (e.g., breakfast food). As they gain increased semantic knowledge, they shift to more adultlike taxonomic (e.g., food)-based strategies for categorization (Nelson & Nelson, 1990). Item typicality influences the likelihood that an item is retrieved and produced. A factor interacting with typicality and influencing categorization and retrieval of items is the frequency of use of the category item.

The category fluency task is composed of three main processes engaging the semantic knowledge and executive function to varying degrees (Raboutet et al., 2010). From a spreading activation account (Collins & Loftus, 1975), the category fluency task engages a search process through the semantic network called *hierarchical exploration*. Participants employ an intracategorical process called *clustering* (Raboutet et al., 2010) to generate a sequence of semantically similar exemplars within the category until the item options are exhausted. Within the network, they generate subsets of related items (e.g., “dog” and “cat”) and then

shift to another related subcategory of exemplars (e.g., “lion” and “tiger”). The latter process is termed as *intercategorical* or *switching* (Raboutet et al., 2010). Success in this task largely depends on the semantic information stored for each category exemplar.

Research using category fluency tasks has examined the number of items produced, the number of errors, and the type of errors as a function of time (Crowe, 1998; Hurks et al., 2006; Raboutet et al., 2010). Participants produce more items during the initial period (e.g., the first 15–30 s) compared to the later part of the task (the final 30 s of a 60-s task). During the initial period, automatic information processing facilitates retrieval, and as time elapses, retrieval is then dominated by controlled information processing (Hurks et al., 2006). High-frequency typical exemplars are more easily accessible (Rosch, 1973) and are thus more likely to be produced during the initial period during automatic processing. In contrast, low-frequency, atypical exemplars occur mostly during the later time course and involve controlled processing (Crowe, 1998; Hurks et al., 2006). Perseverative errors are more likely during the second half of the period (Raboutet et al., 2010).

Category fluency has also been investigated in bilingual speakers to examine the influence of dual language exposure on such tasks. The experience of interacting regularly in two languages has been documented to strengthen speakers’ verbal working memory and executive control in tasks such as inhibiting distractions, conflict resolution, selective attention, and switching (Bialystok, 2005, 2007, 2015; Blom, Küntay, Messer, Verhagen, & Leseman, 2014; Carlson & Meltzoff, 2008; Poulin-Dubois, Blaye, Coutya, & Bialystok, 2011). These enhanced skills may positively influence bilingual speakers’ performance in category fluency tasks. For example, stronger working memory might help speakers monitor their output and avoid repetition. In contrast, by the very nature of operating in two languages, bilinguals have less experience in each of their two languages, leading to reduced lexical knowledge in each language individually. In category fluency tasks, which depend heavily on linguistic representation and lexical retrieval, bilingual participants tend to perform more poorly than their monolingual counterparts (Gollan, Montoya, & Werner, 2002; Rosselli et al., 2002).

Bilingual exposure exerts considerable influence on the structures of lexical categories of the two languages. Ameel, Malt, Storms, and Van Assche (2009) studied category structures using typicality ratings in Dutch–French bilingual and monolingual adult speakers. They found simplified category boundaries in bilinguals. Atypical exemplars at the category boundaries had fewer language-specific features in bilingual compared to monolingual speakers. They concluded that bilingual language experience affects the category structure and differences were more pronounced for the atypical items occurring at the boundaries of categories.

From a theoretical perspective, the representation of low-frequency and atypical category members can draw partial support from the distributed feature model (Van

Hell & De Groot, 1998) of bilingual semantic representation. The model assumes that feature knowledge is shared across the two languages of bilinguals. The category members that are typical only to a language and culture share lesser overlap of conceptual features. Also, a word in the lexicon can be represented as a pattern of activation across a network of interconnected units or semantic features (Masson, 1995). Repeated exposure to these co-occurring activations of a set of semantic features strengthens associative learning, also called as *Hebbian learning* (Hebb, 1949). With respect to atypical items of the category, they share fewer features with the category prototype and also occur less frequently in the linguistic environment. These features may also be idiosyncratic and occur very rarely in the input of individuals influencing their representation.

Bilingual children also tend to produce different exemplars for categories in each of their language, reflecting their distributed language experience. However, in conceptual scoring of the category items in their two languages, with translational equivalents counted only once, their performance tends to be comparable to that of monolinguals (Peña, Bedore, & Zlatic-Giunta, 2002). Bilingual children also appear to shift to adultlike taxonomic categorization strategies at earlier ages compared to monolingual children (Peña et al., 2002; Sheng & Lam, 2015). Thus, it is thought that the process of having to learn a greater number of words across their two languages may push children into using a more adultlike taxonomic strategy earlier. Recently, Shivabasappa, Peña, and Bedore (2017) evaluated the effects of language use and age on the generation of typical category members in a semantic category fluency in Spanish–English bilingual children and adults. They determined the typicality of category items based on adult production frequency, with typical items having higher production frequencies. Both older and younger children produced a similar number of typical items. However, older children produced typical items earlier in their word lists. These results were influenced by condition and language. Children produced typical items earlier in their list in the slot-filler condition and when tested in English. Thus, bilingual children are sensitive to input frequencies of each of their languages and appear to converge on typical items. With development, children access typical items earlier in their word lists, demonstrating a strengthening in the consolidation of typical items.

Convergence in Category Fluency Tasks

Studying semantic convergence patterns helps us understand at what point along the semantic fluency task bilingual children with DLD start to differ from TD children and also if there are any interactions between reduced frequency in input language and limited semantic knowledge leading to the divergence. Examining semantic convergence in children's slot-filler and taxonomic categories helps to further understand the categorization strategies children use in a bilingual context. Slot-filler categories rely heavily on contextual or thematic uses, whereas taxonomic categorization

involves a higher level of hierarchical relations that are not readily available in the immediate linguistic context. Studying convergence in the use of these two strategies with adult patterns may provide greater insight about how bilingual children transition from context-based to rule-based taxonomic categories in their two languages.

In this study, we compared convergence/divergence of semantic category items in Spanish–English bilingual children with and without DLD with reference to a bilingual adult norm to address the following research questions:

1. In a category generation task, at what position in the list do children with DLD diverge from their TD peers in their English and Spanish semantic categories?
2. How much does the convergence of category items in bilingual children with and without DLD vary with respect to slot-filler versus taxonomic conditions in their English and Spanish responses?

Based on the previous studies, we hypothesized that children with DLD may tend to diverge at later positions in the category fluency tasks that require greater semantic knowledge and more controlled processing (Crowe, 1998; Hurks et al., 2006; Raboutet et al., 2010). With respect to the slot-filler/taxonomic condition, we predicted greater convergence in the slot-filler condition, as it is an early developing categorization strategy (Nelson & Nelson, 1990; Peña et al., 2002; Sheng & Lam, 2015).

Method

Participants

The participants were 37 Spanish–English bilingual children with DLD and 37 Spanish–English bilingual TD children. They were part of a larger study ($n = 186$) of semantic and syntactic development of Spanish–English bilingual children (Peña, Bedore, & Fiestas, 2013; Sheng et al., 2012). Children in the current study are all those with DLD and their age-, language-, and exposure-matched TD peers. The demographic details of the participants are provided in Table 1. Children's chronological age ranged from 7;0 to 9;11 (years;months; $M = 8;26$, $SD = 0.96$). All children were heritage Spanish speakers, identified as Hispanic (per parent report), and lived in Texas or Colorado. Besides the child participants, there were 20 Spanish–English bilingual adult participants ($M_{age} = 21.35$, $SD = 2.18$; see Table 1). They were heritage Spanish speakers living in central Texas. The adult participants and TD children were part of the data set reported in the study by Shivabasappa et al. (2017).

Parents provided an hourly account of languages the child heard and spoke on a typical weekday and weekend day. Similarly, teachers completed a questionnaire of each child's English and Spanish use during every school hour. The data from parent and teacher reports were combined to calculate a weighted average of current Spanish and English input and output. Age of first exposure to English was determined by the parent report of Spanish and/or English use on a year-by-year basis.

Table 1. Participant details.

Variable	Children with DLD (37)	Children with TD (37)	Adults (20)
Age (years)	8.33 (0.99)	8.33 (0.98)	21.35 (2.18)
Current English use (%)	43 (13)	43 (12)	72 (14)
Age of English onset (years)	4.08 (1.86)	3.95 (1.72)	4.05 (2.01)
SES ^a	20.91 (8.82)	21.32 (8.96)	NA
Education ^b (years)	NA	NA	16.23 (1.52)
English proficiency ^c	2.81 (0.88)	3.86 (0.84)	4.89 (0.18)
Spanish proficiency ^c	3.63 (0.68)	4.42 (0.95)	4.16 (0.50)

Note. DLD = developmental language disorder; TD = typical development; NA = not applicable.

^aSocioeconomic status (SES) scores were calculated using the Hollingshead (1975) Four-Factor Index of Social Status.

^bThree adults did not report number of years of education. ^cLanguage proficiency was rated using a 5-point scale (children's ratings were obtained from combined ratings of parents and teachers).

The amount of current English and Spanish use in both groups ranged from 18% to 80% ($M = 43\%$, $SD = 14$) and 20% to 80% ($M = 57\%$, $SD = 14$), respectively. The children in the TD and DLD groups were matched pairwise with respect to chronological age (within 4 months), the amount of current English and Spanish use (within 12%), and age of first English exposure (within 1 year: 34 pairs; within 2 years: one pair; within 3 years: one pair). In addition, parents and teachers rated the language skills of children on a 5-point rating scale (1 = *low proficiency*, 5 = *high proficiency*) in three domains, namely, comprehension, vocabulary, and grammar, in both languages. They also described concerns, if any, regarding the child's language abilities. Children's language samples collected from three short narratives in Spanish and English were analyzed for grammaticality. The narratives were transcribed and coded for grammaticality by bilingual research assistants. Next, the proportions of grammatical utterances were calculated using Systematic Analysis of Language Transcripts (Miller & Iglesias, 2010).

The criteria for grouping a child as having DLD were as follows: mean parent and teacher rating of more than 1 SD below the participant pool's mean, valid speech and language concerns, < 80% grammaticality in language samples in both languages, and enrollment in speech and language services. Children who met at least three criteria out of four were identified with DLD. Thirty-seven children (of 280 tested) met the first three criteria. Of these, 35 were currently enrolled in speech and language services with a prior diagnosis of language impairment.

The adult participants completed a language history and current language use questionnaire (Kiran, Peña, Bedore, & Sheng, 2010). Participants reported typical speech and language development and no history of communication impairments. Their current usage ranged from 42% to 95% for English and from 5% to 58% for Spanish.

Procedure

All participants completed a category fluency task of naming items belonging to the three taxonomic semantic categories, namely, animals, food, and clothes, and four

slot-filler categories, namely, farm animals, zoo animals, lunch foods, and winter clothes, in English and Spanish. These semantic categories were originally derived from Nelson and Nelson (1990) and, using a dual-focus approach, were translated into Spanish (Erkut, Alarcón, Coll, Tropp, & García, 1999). Specifically, in the dual-focus approach, the elicitation prompts were developed in both languages, translated, and piloted to retain the prompts that elicited the most robust responses in both languages. We used item analysis of responses from approximately 700 children who participated in the Bilingual English–Spanish Assessment (BESA; Peña, Gutierrez-Clellen, Iglesias, Goldstein, & Bedore, 2014) normative study to reduce the number of semantic categories to the current set. Hence, out of three slot-filler categories for each taxonomic category, the seven that elicited the most responses from TD children across both Spanish and English were retained.

Testing for category fluency was conducted individually for both child and adult participants, in the context of other semantic tasks such as repeated associations, definitions, comparisons, and description. In the category fluency task, participants named as many items as possible belonging to the target category within a time limit of 60 s. Participants listed items for each of the seven categories in both English and Spanish. Testing in each language was completed in two separate blocks over three to four sessions while counterbalancing order of language first tested. The orders of presentation of semantic categories in each language were also varied taking into account that the taxonomic and slot-filler items of same semantic categories (e.g., animals and farm animals) are not presented consecutively.

Children were tested in the schools in a quiet area. They were given the instructions “Tell me all the _____ (e.g., animals) you can think of. Ready? Start” and allowed 60 s to respond. They were prompted (“And ...?” “Is there more?”) or given back-channeling cues (e.g., “uh-huh,” “mmm-hmmm”) to continue responding if they paused. Responses were audio-recorded while, simultaneously, the examiners transcribed responses during testing. Audio-recorded responses were cross-checked for the accuracy of the transcribed responses. Repeated and code-switched items were excluded from the analysis. The code-switched

items were excluded in order to examine category fluency in each of the test languages separately.

The adult participants were tested in the lab in two sessions of two blocks each. In a given session, one block was presented in one language, and then the next block was presented in the other language. Semantic categories were counterbalanced across blocks and sessions so that the same categories were not given in both languages in the same day. Language used in each block was counterbalanced over participants. Similar to the child participants, the same taxonomic and slot-filler categories were never presented consecutively. All responses were audio-recorded and transcribed.

Study Measures

Preliminary analysis of transcribed responses showed that children with DLD had more errors ($M = 0.98$, $SD = 1.48$) compared to TD children ($M = 0.61$, $SD = 0.98$). Children with DLD also produced more responses in the non-test language ($M = 1.39$, $SD = 2.28$) compared to TD children ($M = 0.50$, $SD = 1.07$). Errors, code-switched items, and repeated responses were excluded from further analysis.

This study aimed to examine how closely category items generated by children with and without DLD converged with adult-generated items. As a first step, the data collected from adult participants were analyzed to tabulate the items listed for the seven semantic categories. Next, for each item of a category, we counted the number of adult participants out of 20 who had listed the item (production frequency). Based on this adult production frequency, we identified 10 items for each category with the highest production frequencies (the lists of 10 items in English and Spanish semantic categories and their production frequencies are reported in Supplemental Material S1). Preliminary examination of the top 10 items from the adult participants revealed that, out of 140 words from seven categories in each language, 98 items were listed in both languages. Forty of the 98 items (40.82%) were in the top five positions in both languages; 60 of the 98 (61.22%) were in the top five in one or both languages. Of the 42 total items listed in only one language, 10 items (23.81%) occurred in the top five positions.

Each item in the word list was assigned a weight equivalent to the number of adults who had produced that item for that category. For example, if 18 out of 20 adults listed “dog” for the category animal, the item “dog” was assigned a weight of 18. Thus, the weights of each item represent how likely adults are to produce it as the member of the category.

Children’s category fluency data were analyzed for the presence of the identified 10 items in their lists at the top 10 positions. The average number of items generated by children in the DLD and TD groups were as follows: $M = 4.94$, $SD = 3.52$, and $M = 6.88$, $SD = 3.72$, respectively. Hence, we analyzed children’s responses starting from the first three positions to the first 10 positions in their production list. Specifically, we calculated a measure of convergence scores for top three to top 10 positions based on the items

generated and their respective weights. For instance, while analyzing for items in top three positions for a category, if a child listed two out of the top three identified items (from an adult’s list) with weights of 18 and 17, respectively, the child received a convergence score of $18 + 17 + 0 = 35$ for that category. Higher convergence scores indicated more congruence with adult category structure. Next, we calculated convergence scores for the four items generated by the child in the first four or the top four positions. Thus, we calculated the convergence scores for the top three through top 10 items across all seven categories in both languages.

Data Analysis

The first research question addresses the effect of group and semantic categories on the convergence/divergence of children in English and Spanish. To answer this question, the scores were calculated by combining the four slot-filler categories with their respective taxonomic categories, namely, animals, foods, and clothes, in both test languages. For the second research question addressing the effect of taxonomic and slot-filler conditions on convergence scores, the scores were calculated by grouping the categories animals, foods, and clothes under the taxonomic condition and the categories zoo animals, farm animals, winter clothes, and lunch foods under the slot-filler condition in both the test languages. The convergence scores obtained from children’s data were subjected to statistical analyses using mixed-model analyses of covariance (ANCOVAs). The p values were adjusted for multiple comparisons using a Benjamini–Hochberg correction. Effect sizes (η_p^2) are interpreted based on Cohen’s (1988) guidelines where small = .009, medium = .059, and large = .138.

Results

Group, Semantic Categories, and Test Language Comparisons

We were first interested in whether there were differences by group and semantic category in the convergence of the top three to top 10 items generated by the children. The convergence scores obtained from children were indicative of how closely their semantic categories resembled the adult semantic categories. Higher scores indicated greater convergence with adult category structures based on adult production frequencies, as described in the Data Analysis section. To address the first research question, the four slot-filler categories were collapsed with their respective three taxonomic categories (animals, clothes, and food). The mean convergence scores across semantic categories of TD children and children with DLD are reported in Table 2. For these data, we further conducted eight mixed ANCOVAs with the between-subjects factor of group (DLD and TD) and within-subject factors of category (animals, food, and clothes) and test language (Spanish and English), with the total number of correct items as the covariate. The convergence scores for the top three through top 10 items were the

Table 2. Mean (adjusted) convergence scores (*SE*) across semantic categories of typically developing (TD) children and children with developmental language disorder (DLD).

Items	Groups	All categories	Animals	Clothes	Food
Top 3	TD	19.59 (0.59)	26.22 (0.90)	20.59 (1.06)	8.67 (1.13)
	DLD	18.44 (0.63)	22.24 (0.87)	20.69 (1.40)	10.52 (1.10)
Top 4	TD	23.99 (0.66)	31.88 (1.03)	25.11 (1.21)	11.06 (1.28)
	DLD	22.65 (0.71)	26.85 (0.99)	25.75 (1.59)	13.28 (1.25)
Top 5	TD	28.52 (0.72)	36.44 (1.12)	31.30 (1.31)	13.89 (1.39)
	DLD	27.21 (0.77)	31.83 (1.07)	31.43 (1.72)	16.08 (1.36)
Top 6	TD	32.01 (0.75)	41.01 (1.15)	35.26 (1.35)	15.27 (1.44)
	DLD	29.96 (0.80)	35.18 (1.11)	34.69 (1.78)	17.43 (1.40)
Top 7	TD	35.09 (0.76)	44.00 (1.18)	39.71 (1.39)	17.11 (1.47)
	DLD	33.14 (0.82)	37.99 (1.14)	40.12 (1.83)	18.92 (1.44)
Top 8	TD	37.25 (0.78)	46.10 (1.20)	43.05 (1.41)	18.18 (1.50)
	DLD	35.33 (0.83)	40.49 (1.16)	43.09 (1.86)	19.84 (1.46)
Top 9	TD	39.34 (0.78)	48.23 (1.20)	45.67 (1.41)	19.69 (1.50)
	DLD	37.17 (0.83)	42.38 (1.16)	45.73 (1.86)	20.81 (1.46)
Top 10	TD	40.97 (0.77)	50.34 (1.20)	47.38 (1.41)	20.52 (1.49)
	DLD	38.96 (0.83)	44.79 (1.15)	47.59 (1.85)	21.61 (1.46)

dependent measures. Hence, there were eight dependent measures, one for each of the top three to top 10 items. The results of the analyses are reported in Table 3.

The results of these ANCOVAs showed that there were no significant main effects for group for the top three items (DLD: $M = 18.44$, $SE = 0.63$; TD: $M = 19.59$, $SE = 0.59$), top four items (DLD: $M = 22.65$, $SE = 0.71$; TD: $M = 23.99$, $SE = 0.66$), or top five items (DLD: $M = 27.21$, $SE = 0.77$; TD: $M = 28.52$, $SE = 0.72$). The mean convergence scores across the top three to top 10 positions in children with and without DLD are shown in Figure 1. This indicates that children with DLD and TD children produced similar items in the top three to top five positions, after controlling for the total number of correct responses. However, there were significant main effects for top six (DLD: $M = 29.96$, $SE = 0.80$; TD: $M = 32.01$, $SE = 0.75$), top seven (DLD: $M = 33.14$, $SE = 0.83$; TD: $M = 35.09$, $SE = 0.76$), top eight (DLD: $M = 33.14$, $SE = 0.82$; TD: $M = 35.03$, $SE = 0.76$), top nine (DLD: $M = 37.17$, $SE = 0.83$; TD: $M = 39.34$, $SE = 0.78$), and top 10 (DLD: $M = 38.96$, $SE = 0.83$; TD: $M = 40.97$, $SE = 0.77$) items (see Figure 1). The

differences between the two groups indicate the items produced from the sixth position onward diverged rather than converged.

We were also interested in whether there were differences by semantic category. Results indicate that there were significant differences for all the analyses from the top three to top 10 items (see Table 3). Children converged with adult norms on category members to a greater extent for the animal category (adjusted mean [SE] values ranging from 24.22 [0.63] to 47.56 [0.83]), followed by clothes (adjusted mean [SE] values ranging from 20.63 [0.88] to 47.48 [1.16]). The smallest convergence was noted for the food category (adjusted mean [SE] values ranging from 9.59 [0.79] to 21.06 [1.04]). With respect to the two test languages, children converged on the category members (as compared to the adult norms) similarly in Spanish (adjusted mean [SE] values ranging from 17.26 [0.57] to 37.70 [0.76]) and English (adjusted mean [SE] values ranging from 20.78 [0.64] to 42.23 [0.85]) across all eight comparisons.

There were significant interactions between group and category for all the eight analyses (see Table 3). Post

Table 3. Results of analyses of covariance for semantic categories (adjusted p values).

Items	Language ability ($df = 1$, $df\ error = 72$)			Category ($df = 2$, $df\ error = 72$)			Test language ($df = 1$, $df\ error = 72$)			Language Ability $\times$ Category ($df = 2$, $df\ error = 72$)		
	F	p^a	η_p^2	F	p^a	η_p^2	F	p^a	η_p^2	F	p^a	η_p^2
Top 3	2.80	.09	.002	126.70	< .001*	.18	5.11	.21	.01	5.48	.007*	.010
Top 4	3.37	.08	.002	133.33	< .001*	.19	0.63	.42	.004	6.59	.005*	.012
Top 5	3.61	.08	.001	146.77	< .001*	.21	0.68	.42	.005	4.35	.013*	.009
Top 6	5.56	.04*	.003	178.65	< .001*	.24	0.79	.42	.006	5.98	.005*	.011
Top 7	5.38	.04*	.003	198.95	< .001*	.27	1.89	.27	.009	6.06	.005*	.011
Top 8	5.38	.04*	.003	198.95	< .001*	.27	1.89	.27	.009	6.06	.005*	.011
Top 9	6.35	.04*	.003	237.42	< .001*	.30	3.63	.24	.014	4.97	.008*	.009
Top 10	5.00	.04*	.003	268.31	< .001*	.33	2.85	.25	.013	5.14	.008*	.009

Note. The p values are adjusted using Benjamini–Hochberg correction for multiple comparisons.

^aAdjusted p values.

* $p < .05$.

Figure 1. Mean convergence scores in the top three to top 10 positions in children with and without developmental language disorder (DLD). TD = typically developing.

hoc comparison using Tukey's method revealed that children with DLD (adjusted mean [SE] values ranging from 22.23 [0.87] to 44.78 [1.15]) produced significantly lower scores than their TD peers (adjusted mean [SE] values ranging from 26.21 [0.90] to 50.33 [1.19]) for the category of animal but did not differ for the category of clothes (DLD: adjusted mean [SE] values ranging from 20.69 [1.40] to 47.59 [1.85]; TD: adjusted mean [SE] values ranging from 20.58 [1.06] to 47.37 [1.40]) and food (DLD: adjusted mean [SE] values ranging from 10.52 [1.10] to 21.61 [1.45]; TD: adjusted mean [SE] values ranging from 8.67 [1.13] to 20.51 [1.49]). Both groups of children showed significantly lower convergence scores for the category of food compared to both animals and clothes. However, they scored similarly for the animals and clothes categories. There were no other significant interactions.

Taxonomic/Slot-Filler Condition, Group, and Test Language Comparisons

With the aim of understanding the effects of taxonomic and slot-filler conditions on category convergence, the three categories were regrouped by these two conditions for the next set of analyses. Hence, there were three categories in taxonomic condition (animals, clothes, and food) and four in slot-filler (zoo animals, farm animals, winter clothes, and lunch food). The mean convergence scores across the slot-filler and taxonomic conditions of TD children and children with DLD are reported in Table 4.

Similar to the previous section, mixed-model ANCOVAs were used with the between-subjects factor group (DLD and TD) and the within-subject factors condition (taxonomic and slot-filler) and test language (Spanish and English). Similar to previous analyses, the convergence scores for the top three through top 10 items were the dependent measures. The number of correct items was the covariate in the models. The results of the analyses are reported in Table 5. The results showed a significant main effect for condition

Table 4. Mean (adjusted) convergence scores (SE) across slot-filler and taxonomic conditions of typically developing (TD) children and children with developmental language disorder (DLD).

Items	Groups	Slot-filler condition	Taxonomic condition
Top 3	TD	21.54 (0.81)	16.47 (1.08)
	DLD	21.81 (0.94)	15.61 (0.98)
Top 4	TD	27.40 (0.90)	18.64 (1.20)
	DLD	28.31 (1.05)	17.36 (1.09)
Top 5	TD	32.16 (0.99)	23.24 (1.31)
	DLD	33.01 (1.15)	21.43 (1.20)
Top 6	TD	36.14 (1.04)	26.13 (1.38)
	DLD	36.46 (1.20)	24.01 (1.25)
Top 7	TD	39.79 (1.08)	28.47 (1.43)
	DLD	39.68 (1.25)	27.07 (1.30)
Top 8	TD	42.09 (1.12)	30.39 (1.48)
	DLD	42.11 (1.29)	29.04 (1.35)
Top 9	TD	44.22 (1.13)	32.11 (1.51)
	DLD	43.94 (1.31)	30.64 (1.37)
Top 10	TD	45.55 (1.16)	34.30 (1.55)
	DLD	45.50 (1.35)	32.99 (1.41)

for all the eight analyses. Children had significantly higher convergence scores for slot-filler categories (mean values ranging from 21.67 [0.62] to 45.52 [0.88]) compared to taxonomic categories (mean values ranging from 16.03 [0.72] to 33.64 [1.04]). The mean convergence scores across the taxonomic and slot-filler conditions are shown in Figure 2. There were no significant main effects for group or test language (see Table 5).

There were significant interactions between condition and test language for all the eight analyses (see Table 5). Post hoc comparison using Tukey's method showed that convergence scores for the slot-filler condition were higher in English (adjusted mean [SE] values ranging from 24.66 [0.92] to 50.65 [1.32]) than in Spanish (adjusted mean [SE] values ranging from 18.68 [0.83] to 40.38 [1.19]; $p < .001$). However, the scores did not differ for the taxonomic condition in English (adjusted mean [SE] values ranging from 15.85 [1.01] to 31.51 [1.45]) and Spanish (adjusted mean [SE] values ranging from 16.21 [1.04] to 35.78 [1.50]; $p > .05$). The convergence scores were also higher in the English slot-filler condition (adjusted mean [SE] values ranging from 24.66 [0.92] to 50.65 [1.32]) compared to the English taxonomic condition (adjusted mean [SE] values ranging from 15.85 [1.01] to 31.51 [1.45]; $p < .001$) and the Spanish taxonomic condition (adjusted mean [SE] values ranging from 16.21 [1.04] to 35.78 [1.50]; $p < .001$). Children also produced higher scores in the Spanish slot-filler condition (adjusted mean [SE] values ranging from 18.68 [0.83] to 40.38 [1.19]) compared to the Spanish taxonomic (adjusted mean [SE] values ranging from 16.21 [1.04] to 35.7 [1.50]; $p < .001$) and English taxonomic (adjusted mean [SE] values ranging from 15.85 [1.01] to 31.51 [1.45]; $p < .001$) conditions.

Discussion

This study examined the amount of convergence of semantic category items using a category fluency task in bilingual children with and without DLD compared to bilingual

Table 5. Results of analyses of covariance for taxonomic and slot-filler conditions.

Items	Language ability (<i>df</i> = 1, <i>df</i> error = 72)			Condition (<i>df</i> = 1, <i>df</i> error = 72)			Test language (<i>df</i> = 1, <i>df</i> error = 72)			Condition × Test Language (<i>df</i> = 1, <i>df</i> error = 72)		
	<i>F</i>	<i>p</i> ^a	η_p^2	<i>F</i>	<i>p</i> ^a	η_p^2	<i>F</i>	<i>p</i> ^a	η_p^2	<i>F</i>	<i>p</i> ^a	η_p^2
Top 3	2.36	.12	< .001	41.04	< .001*	.03	5.80	.14	.011	9.88	.003*	.010
Top 4	2.78	.11	< .001	94.14	< .001*	.08	0.91	.34	.004	4.64	.037*	.005
Top 5	2.87	.11	< .001	91.82	< .001*	.07	1.03	.34	.005	0.12	.72	.000
Top 6	4.51	.07	.001	95.50	< .001*	.08	1.21	.34	.005	7.06	.013*	.007
Top 7	4.47	.07	.001	100.10	< .001*	.08	2.22	.22	.007	18.24	< .001*	.018
Top 8	4.47	.07	.001	100.10	< .001*	.08	2.22	.22	.007	18.24	< .001*	.018
Top 9	5.20	.07	< .001	98.98	< .001*	.08	3.80	.21	.010	18.70	< .001*	.018
Top 10	3.93	.08	< .001	82.80	< .001*	.07	2.79	.22	.018	26.41	< .001*	.025

Note. *p* Values are adjusted using Benjamini–Hochberg correction for multiple comparisons.

^aAdjusted *p* values.

**p* < .05.

adult norms. We also studied how the extent of convergence may vary with respect to the different categories and the two test languages and in taxonomic versus slot-filler conditions. Using a qualitative approach, we studied category exemplars generated by children as a function of its position in the sequence produced over time. For this purpose, every item produced by adults for each category in each test language was assigned a weight equivalent to adults' production frequency. Typical items have higher weights as adults list them more often. For each child, we calculated a convergence score by adding the weights (derived from adult production frequency) of all the items the child had produced. We further examined how convergence scores vary with each position over time in a category fluency task.

Group, Semantic Categories, and Test Language Comparisons

Group

Results revealed that children with DLD did not differ significantly from TD children on convergence scores for items in the top three to top five positions across the

categories tested. Both groups show similar convergence with adults for category exemplars produced at the earlier positions during the task. However, for items produced in the top six to top 10 positions, children with DLD showed increased divergence from their peers. The results were similar to those reported by Raboutet et al. (2010) for time course analysis of category fluency output. In their study, participants produced more items in the initial period that corresponds to earlier positions in our study. Similar to Crowe (1998), items produced earlier in the task were typical exemplars of the category. Access of typical and high-frequency items is less effortful and automatic (Hurks et al., 2006). These items are very easily activated because of the stronger representation resulting from robust connections, which strengthen with frequent retrieval (Griffin & Bock, 1998). Here, we observed that children with DLD were as likely as their TD peers to retrieve typical items.

The results reflecting similar convergence for top five items can also derive support from the distributed feature model (Van Hell & De Groot, 1998) of bilingual semantic representation. During the preliminary analysis of the adult-generated word lists (see Study Measures section), it was noted that items occurring in both languages more often occurred in the top five positions. This indicates shared conceptual knowledge that would further strengthen their semantic representations from repeated exposures and activation in both languages.

In the later positions of the task, scores of children with DLD remained lower, indicating that they were at a loss to generate new convergent items. TD children, however, showed relatively higher scores as we analyzed items in the later positions. These results can be attributed to interaction of three factors, namely, lexical–semantic deficits in DLD, nature of bilingual exposure, and properties of atypical items, as described in the following paragraphs.

Children with DLD have less robust lexical semantic networks, which impede the easy acquisition of new related words of the category (Beckage, Smith, & Hills, 2011). The deficits in vocabulary, along with reduced verbal working

Figure 2. Mean convergence scores for the taxonomic and slot-filler conditions in the top three to top 10 positions.

memory, may have contributed to additional difficulties in children with DLD at later positions, resulting in reduced convergence scores in these positions. It is also true that category fluency is highly dependent on lexical–semantic knowledge rather than executive control. Bilingual exposure enhances executive control in individuals but simultaneously puts them at disadvantage for lexical–semantic knowledge in each language. Category fluency is vulnerable in children with DLD where the enhanced executive control, if any, cannot offset their deficits in lexical–semantics. Sheng et al. (2013) also reported similar results of lower convergence in bilingual children with DLD in a repeated word association task.

Furthermore, the memberships of atypical members are not based on similarity to the prototype, are acquired through exposure to their name–object pairings, and hence are more difficult to learn through generalization (Ameel et al., 2009). Atypical items comprising mostly idiosyncratic and fewer shared features co-occur in lower frequencies in the input hindering their learning in children with DLD. Children with DLD present with greater difficulties with low-frequency and atypical category items, indicating that they require a greater amount of language experience to assimilate and consolidate their learning. In this study, adult participants showed less agreement for the items produced at later positions of the task, indicating increased variability in terms of occurrence and category membership (see Supplemental Material S1). Children with DLD thus need more exposures for lexical learning, making less frequent targets more difficult to retain. This effect may reflect in their lower convergence scores for the items generated in the later positions compared to their TD peers.

Semantic Categories

With respect to categories, the convergence scores were significantly different across three categories, with animals showing the highest convergence scores followed by clothing. The category of food had the least convergence scores. Similar results were seen in previous studies involving these categories (Ross & Murphy, 1999; Shivabasappa et al., 2017). This pattern of convergence reflects the nature of these semantic categories. The items of some semantic categories, namely, animals, have a clear set of inclusionary criteria and properties the exemplars should possess. The semantic properties and information about these members are almost similar across the population, as individuals rarely interact with them on a day-to-day basis. These concepts are built upon experiences that are comparable among individuals, as they may gain exposure to them through books, television programs, and academic settings. Categories such as food, on the other hand, have fairly clear boundaries, but individual family and cultural influences create a progressive scale to which items adhere to categories (Ross & Murphy, 1999). Children with DLD scored lower than TD children for the category of animal compared to food and clothes, indicating lesser integration of semantic knowledge learned through language experience.

Test Language

Children in both groups showed a similar pattern of convergence scores across their two test languages. Children with DLD, despite lower convergence scores resulting from their shorter list of responses, performed similar to TD peers in the two test languages, with similar convergence scores in both of their languages. Categorization is a basic skill acquired early in age, and children are able to transfer their categorization ability from one language to another. This is manifested in this study as children produced a similar number of typical items in both test languages. However, all the items that they tend to generate in each category for each language may not be similar to the items generated by their monolingual peers, as the category structure may vary across languages due to the differences in the way how concepts are expressed at the linguistic level (Ameel et al., 2009).

The results further support the use of category fluency tasks for the assessment and screening of bilingual children with DLD, as children with DLD perform poorly in the later positions of the task. In addition, bilingual children tend to perform comparably in both of their languages, as categorization is a basic, transferable cognitive skill. Using semantic categories of animals and clothing as probes is more suitable than the inherently less converging category of food.

Taxonomic/Slot-Filler Condition, Group, and Test Language Comparisons

The convergence scores also varied across the taxonomic and slot-filler conditions. The slot-filler category items showed higher convergence scores than taxonomic categories in both Spanish and English. However, children in both groups performed similarly across these two conditions. Higher convergence scores also indicate that they produced more typical items. Similar findings were also reported by Shivabasappa et al. (2017), who found that TD children produced typical items earlier in the slot-filler conditions than in the taxonomic conditions. Children have more experience categorizing using slot-filler strategies as they begin to use such strategies from younger ages (Nelson & Nelson, 1990). Hence, their slot-filler category structure and items resemble closely to adults', resulting in higher convergence scores. However, categorizing using abstract taxonomic relations requires greater experience with the language. Also, it is interesting that, in the slot-filler condition, children demonstrated higher convergence in English compared to Spanish. Children are exposed to English typically in a classroom setting, and they get exposed to the community with a majority of English speakers. They tend to have a greater amount of uniform experience in English through schooling. This is reflected in their higher convergence scores for exemplars in English. On the other hand, children are typically exposed to Spanish in the home environment with immediate family members and a small community exposure. There can be a lot of room for variability in their language experience (Shivabasappa et al., 2017). This is reflected in their lower convergence scores for Spanish especially in the slot-filler categories.

Limitations and Future Directions

This study, despite providing insights into the nature of categorization skills and convergence patterns, does not permit us to pinpoint the contribution of each of the aspects of semantic deficits such as poor vocabulary, less robust semantic connections, and/or verbal memory limitations and their combined effects toward the categorization and convergence scores in children with DLD. Future studies may focus on teasing apart each of the factors' contributions toward the patterns of deficits in children with DLD. We also know that bilingual language experience is a spectrum ranging from different amounts of input and output and dominance profiles that may change over time in each individual. Hence, future studies may also examine the correlations of the amount of each language use and children's categorization skills in each of their language. Future studies can also explore how semantic fluency tasks may converge or diverge across other language pairs.

Conclusion

To summarize, this study adds to the growing evidence about the sparseness of lexical entries in children with DLD; by studying bilinguals with DLD relative to their TD peers, we are better able to understand how a sparse lexicon contributes to divergence. The weakness or lack of convergence in the children with DLD observed in this study was not surprising; however, the approach we have used is unique, which allows us a more nuanced view of where the breakdowns are with DLD. Children with DLD show similar convergence patterns as their TD peers for the top five earlier produced, frequent, and typical members of the semantic categories. Children's difficulties were more pronounced in the later part of category fluency tasks for items produced after the fifth position. This may reflect their limited knowledge of infrequent and atypical concepts that are usually produced in the later portions of category fluency tasks. Clinically, the results underscore the usefulness of a semantic fluency task for bilingual language screening and assessment in children given that it is a basic skill that converges across languages and children with DLD perform lower on this task. Categories that inherently converge such as animals and clothing would likely be more appropriate than food. It is also crucial that we go beyond typical, frequent category members to support children's development of semantic depth.

Acknowledgments

This research was funded by National Institute of Child Health and Human Development Grant R21HD053223 awarded to Elizabeth D. Peña (PI) and Lisa M. Bedore (Co-PI).

References

- Adams, A. K., & Bullock, D. (1986). Apprenticeship in word use: Social convergence processes in learning categorically related

nouns. In S. A. Kuczaj & M. D. Barrett (Eds.), *The development of word meaning* (pp. 155–197). New York, NY: Springer.

Ameel, E., Malt, B. C., Storms, G., & Van Assche, F. (2009). Semantic convergence in the bilingual lexicon. *Journal of Memory and Language*, 60, 270–290.

Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLOS ONE*, 6(5), e19348.

Bialystok, E. (2005). Consequences of bilingualism for cognitive development. In *Handbook of bilingualism: Psycholinguistic approaches* (pp. 417–432). New York, NY: Oxford University Press.

Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, 10(3), 210–223.

Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, 9(2), 117–121.

Bjorklund, D. F., Thompson, B. E., & Ornstein, P. A. (1983). Developmental trends in children's typicality judgments. *Behavior Research Methods & Instrumentation*, 15(3), 350–356.

Blom, E., Küntay, A. C., Messer, M., Verhagen, J., & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology*, 128, 105–119.

Carlson, S. M., & Meltzoff, A. N. (2008). Bilingual experience and executive functioning in young children. *Developmental Science*, 11(2), 282–298.

Clark, E. V. (2007). Conventionality and contrast in language and language acquisition. *New Directions for Child and Adolescent Development*, 115, 11–23. <https://doi.org/10.1002/cd.179>

Coelho, S., Albuquerque, C. P., & Simões, M. R. (2013). Specific language impairment: A neuropsychological characterization. *Paidéia (Ribeirão Preto)*, 23(54), 31–41.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428.

Crowe, S. F. (1998). Decrease in performance on the verbal fluency test as a function of time: Evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology*, 20(3), 391–401.

Erkut, S., Alarcón, O., Coll, C. G., Tropp, L. R., & García, H. A. V. (1999). The dual-focus approach to creating bilingual measures. *Journal of Cross-Cultural Psychology*, 30(2), 206–218.

Gollan, T. H., Montoya, R. I., & Werner, G. A. (2002). Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology*, 16(4), 562–576.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38(3), 313–338.

Hall, J., McGregor, K. K., & Oleson, J. (2017). Weaknesses in lexical–semantic knowledge among college students with specific learning disabilities: Evidence from a semantic fluency task. *Journal of Speech, Language, and Hearing Research*, 60(3), 640–653.

Hebb, D. O. (1949). The organization of behavior: A neuropsychological theory. *A Wiley Book in Clinical Psychology*, 62–78.

Hedden, T., Lautenschlager, G., & Park, D. C. (2005). Contributions of processing ability and knowledge to verbal memory tasks across the adult life-span. *The Quarterly Journal of Experimental Psychology Section A*, 58(1), 169–190.

Hollingshead, A. B. (1975). *Four factor index of social status*. Unpublished manuscript, Yale University, New Haven, CT.

- Hughes, D. L., & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical and Experimental Neuropsychology*, 24(5), 642–654.
- Hurks, P. P., Schrans, D., Meijs, C., Wassenberg, R., Feron, F. J. M., & Jolles, J. (2010). Developmental changes in semantic verbal fluency: Analyses of word productivity as a function of time, clustering, and switching. *Child Neuropsychology*, 16(4), 366–387.
- Hurks, P. P. M., Vles, J. S. H., Hendriksen, J. G. M., Kalff, A. C., Feron, F. J. M., Kroes, M. (2006). Semantic category fluency versus initial letter fluency over 60 seconds as a measure of automatic and controlled processing in healthy school-aged children. *Journal of Clinical and Experimental Neuropsychology*, 28, 684–695.
- Kiran, S., Peña, E., Bedore, L., & Sheng, L. (2010). *Evaluating the relationship between category generation and language use and proficiency*. Paper presented at the Donostia Workshop on Neurobilingualism, San Sebastian, Spain.
- Kohnert, K., & Kan, P. (2007). Lexical skills in young children learning a second language: Methods, results and clinical applications. In J. Centeno, R. Anderson, & L. Obler (Eds.), *Communication disorders in Spanish speakers: Theoretical, research and clinical aspects* (pp. 156–168). Tonawanda, NY: Multilingual Matters Ltd.
- Masson, M. E. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(1), 3–23.
- Miller, J., & Iglesias, A. (2010). Systematic Analysis of Language Transcripts (SALT) [Research Version 2010]. Madison, WI: SALT Software.
- Nelson, K., & Nelson, A. P. (1990). Category production in response to script and category cues by kindergarten and second-grade children. *Journal of Applied Developmental Psychology*, 11(4), 431–446.
- Peña, E. D., Bedore, L. M., & Fiestas, C. (2013). Development of bilingual semantic norms: Can two be one. In V. C. M. Gathercole (Ed.), *Solutions for the assessment of bilinguals* (pp. 103–124). Bristol, United Kingdom: Multilingual Matters.
- Peña, E. D., Bedore, L. M., & Zlatić-Giunta, R. (2002). Category-generation performance of bilingual children: The influence of condition, category, and language. *Journal of Speech, Language, and Hearing Research*, 45(5), 938–947.
- Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2014). *Bilingual English–Spanish Assessment (BESA)*. San Diego, CA: AR-Clinical.
- Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology*, 108(3), 567–579.
- Raboulet, C., Sauzéon, H., Corsini, M. M., Rodrigues, J., Langevin, S., & N'Kaoua, B. (2010). Performance on a semantic verbal fluency task across time: Dissociation between clustering, switching, and categorical exploitation processes. *Journal of Clinical and Experimental Neuropsychology*, 32(3), 268–280.
- Restrepo, M. A., & Kruth, K. (2000). Grammatical characteristics of a Spanish–English bilingual child with specific language impairment. *Communication Disorders Quarterly*, 21(2), 66–76.
- Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language*, 71(2), 267–284.
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, 4(3), 328–350.
- Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology*, 38(4), 495–553.
- Rosselli, M., Ardila, A., Salvatierra, J., Marquez, M., Luis, M., & Weekes, V. A. (2002). A cross-linguistic comparison of verbal fluency tests. *International Journal of Neuroscience*, 112(6), 759–776.
- Sheng, L., Bedore, L. M., Peña, E. D., & Taliancich-Klinger, C. (2013). Semantic convergence in Spanish–English bilingual children with primary language impairment. *Journal of Speech, Language, and Hearing Research*, 56(2), 766–777.
- Sheng, L., & Lam, B. P. (2015). Slot-filler and taxonomic organization: The role of contextual experience and maternal education. *Journal of Educational and Developmental Psychology*, 5(1), 128.
- Sheng, L., Peña, E. D., Bedore, L. M., & Fiestas, C. E. (2012). Semantic deficits in Spanish–English bilingual children with language impairment. *Journal of Speech, Language, and Hearing Research*, 55, 1–15.
- Shivabasappa, P., Peña, E. D., & Bedore, L. M. (2017). Typicality effect and category structure in Spanish–English bilingual children and adults. *Journal of Speech, Language, and Hearing Research*, 60(6), 1577–1589.
- Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, 50(3), 337–356.
- Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1(3), 193–211.
- Weckerly, J., Wulfeck, B., & Reilly, J. (2001). Verbal fluency deficits in children with specific language impairment: Slow rapid naming or slow to name. *Child Neuropsychology*, 7(3), 142–152.